

How to tell success from failure

Next week's summit will have served its purpose if the participants merely agree that they should meet again. Some popular goals are unattainable, others deserve more attention than they will get.

Two people who have never previously laid eyes on each other will spend Tuesday and Wednesday of next week in each other's company under the most trying circumstances. Of late, President Ronald Reagan and Mr Mikhail Gorbachev have spent much of their energy in mutual public recrimination, which would be understandable against the background of the antagonism of the two governments they respectively lead, but which is hardly calculated to make for a cosy chat at Geneva. To make things worse, they have arranged to meet without knowing just what they will talk about if, as seems likely, they run out of things to say about their prime topic of conversation, arms control. Yet, as if to ensure that their social embarrassment will haunt them for years to come, the two leaders have encouraged people everywhere to expect great things from their encounter. *Pravda*, not to mention the *New York Times*, has for weeks had space for almost nothing else. This is one reason that the meeting is bound to seem a disappointment, a measure of the gap between people's expectations and reality.

Luckily, none of this implies that the meeting will be a waste of everybody's time. Those who expect wonders from Geneva should be content if the two participants agree only that they will meet again within the year. (Exactly a year from now would inconveniently coincide with the mid-term election season in the United States.) That alone would wonderfully repay the cost of all those airline tickets to Geneva. Less than that would be worrying for everybody. So an agreement to meet again is the least that must be looked for. What follows is a check-list of plusses and minuses by which departures of next week's meeting from that minimum should be judged.

Arms control

That arms control should be central to the agenda is natural, although the widespread belief that this is the sole touchstone of success is mistaken. Most of the recrimination of the past few weeks has centred on the US plan to build a defence against ballistic missiles, the Strategic Defense Initiative or SDI. The Soviet Union says that SDI is a provocation, the United States that a programme (not a plan) to make people everywhere safe from the threat of annihilation could not but be a boon, and that in any case the programme consists only of research; it could become a plan only a few presidential terms from now. Both sides are on solid ground, as in most confrontations born of conflicting interests. From the visit to Moscow earlier this month of Mr George Schulz, the US Secretary of State, it is clear that there will be little give or take on this issue at Geneva. The worst case will be that in which one or other of the participants digs in his heels on SDI, so as provoke a summit meeting as indistinguishable from a shouting match as that in Vienna in 1961, when the late Mr Nikita Khrushchev could not control his temper, read the still-new President John F. Kennedy a demeaning lesson and laid the ground for the Cuban missile crisis almost exactly a year later. The best hope is that next week's actors will perform a little more imaginatively.

A few ingredients could realistically be added to the minimum agreement to meet again within the year. In the field of arms control, there is no reason why the two sides should not agree in principle that there should be a treaty to prohibit either of them from destroying the other's Earth satellites. The Soviet Union

has been asking for such a treaty for the best part of a year, but the United States has replied by saying that it is at a disadvantage at this stage in the evolution of these esoteric weapons. Both perceptions may be correct, but the technology of anti-satellite (ASAT) weapons is at such a rudimentary stage that it would be wiser to live with an imbalance than to embark on the development. Thus an agreement in principle (the details would have to be worked out) to negotiate an ASAT treaty could take the edge off next week's disappointment, and would be worth having in itself.

ABM treaty

Another piece of frosting on next week's otherwise disappointing cake would be an agreement, again in principle, that the Anti-Ballistic Missile (ABM) Treaty of 1972, signed when the notion of *détente* seemed set in concrete, should now be reaffirmed and renegotiated. There are two problems. So much SDI talk has seemed to threaten the ABM treaty that reaffirmation is necessary, but the ambiguities that have allowed the US State Department and the Pentagon to disagree about the testing of SDI components need to be resolved. The treaty is ambiguous because its drafters did not know about SDI, but letting it lapse would be a setback. If SDI is indeed only a research programme, reaffirmation would cost nothing. Optimistic realists will be looking for one or other (or, preferably, both) of these refinements in next week's communique. The ground is not nearly well enough prepared for agreement to be reached next week, but a declaration of principle would allow the negotiators to get down to work. But all of this would be extra to the bare minimum agreement.

Familiar components of other arms control agendas will be absent from next week's agreement. Although the Soviet Union has been operating a unilateral moratorium on the testing of nuclear weapons (due to expire at the end of the month), there is no chance that the superpowers will agree next week to move decisively towards a comprehensive test-ban treaty. The United States is at present at a military disadvantage in being lumbered with the MX missile system, and will have a need to test warheads for whatever succeeds as the mainstay of its strategic force; giving up testing now would not be popular in Washington, but both sides are increasingly discomfited by the knowledge that even if there were a test-ban, France and China would stay outside. Difficulties about remote verification also persist, although less seriously than when the last round of serious negotiations began in 1977, but these might well be overcome by more liberal arrangements for *in situ* monitoring and on-site inspection. It would cost Mr Reagan and Mr Gorbachev nothing to agree that the draft treaty of November 1979 should be dusted off to see whether it will suit present circumstances.

But what of the issues central to the negotiations on strategic arms that have been under way at Geneva since the beginning of the year? The high expectations of next week's meeting stem from the enthusiasm that has been endangered for a substantial reduction of nuclear weapons. The two sides have indeed put forward proposals that outwardly differ very little from each other. The general idea is that strategic nuclear forces should be reduced roughly by a half. The Soviet proposal would selectively reduce present nuclear forces, the United States would limit



warheads instead. The obvious difficulty is that the viability of such comprehensive mutual restraint depends crucially on the details of an agreement, not its general outlines. The most that can be expected next week is that the two governments will acknowledge that the gap between them is not huge, and that they will then wish their negotiators at Geneva good luck. That, obviously, would be a plus. The corresponding minus is the danger that the Soviet Union will link the continuation of these potentially fruitful negotiations with a demand that SDI should be abandoned. That could spell a further deterioration of East-West relations, with possibly dreadful consequences. The question would be more rationally discussed when the draft of a treaty on strategic arms limitation exists, which is another argument for a second meeting.

What else?

The upshot is that there can be no firm agreement on arms control next week, but at best a series of agreements to agree. If there is to be a second meeting, the participants should therefore arrange that they will have other things to talk about. But what? Mr Gorbachev has been saying that arms control must be central to next week's talks, but that overlooks the truth that arms control is feasible only when political relationships are secure. President Reagan, trying to broaden next week's agenda, has been advocating (at the UN General Assembly) a partial list of the world's trouble-spots on which both Afghanistan and the Middle East are conspicuous, the second by its absence. In reality, such an agenda would cause more trouble than emollience. In the early 1970s, it proved possible to trade the promise of Soviet friendliness for that of US technology, particularly (as it turned out) the products of US farm technology now accumulating in the silos of the Middle West. No such easy accommodation is now possible. Nor is there much mileage in the almost ritual negotiation of improved scientific collaboration, the standby of politicians in need of something on which to strike agreement.

But why should not the United States and the Soviet Union negotiate some of the practical issues that divide them? The West's embargo on advanced technology, administered by COCOM, would actually be more coherent if the Soviet Union were given a chance to comment on the nonsenses that frequently arise. The endless troubles about spying in the past year could have been avoided if there were a general understanding about the line dividing the tolerable from the unacceptable. And the time will come when the two superpowers must talk about the political framework for the relationships between European states, East and West. It will be an uncovenanted plus next week if, having agreed to meet again, the two strangers agree to broaden future agendas in this way. □

Education in a mess

A new wave of anxiety about US science education may, this time, bring action.

SCIENCE and engineering education in the United States has frequently been held up elsewhere as a shining example to be followed. In Britain, for example, the many studies over the past two years focused on the shortage of scientific and technical skills have ritually applauded what has been done in the United States. But if US science education at the undergraduate level has the appearance of ruddy health, it is only skin deep. The system, if it can be called that, is in trouble. Science education is facing its biggest crisis since the panic that gripped the United States in the early 1960s, after the Soviet Union beat it into space with Sputnik.

It is true that in the United States undergraduate institutions seem to have adapted more quickly than those elsewhere to the booming demand for skills that have anything to do with computers, thanks largely to the recognition by many US science-based businesses of where their interests lie. (It also helps that the tax system is so weighted in favour of corporate giving that it would

be public relations suicide for a major corporation not to be seen to give away a few millions of dollars every year or so.) But there is more to life, and science, than desk-top computers. The problems are many, and are now being spelled out in painful detail to the National Science Board, the policymaking arm of the National Science Foundation. The board, to its credit, has established a committee (inevitably called a task force) to examine the problem and to see what the foundation should do about it. There is mounting and disturbing evidence that the quality of teaching in both public and private universities is declining. Faculty are ageing and are not being replaced at a sufficient rate, especially in engineering. Laboratory instrumentation is, despite corporate munificence, seriously out of date, particularly in institutions not famed as research establishments. And increasing teaching loads all too often force universities and colleges to rely on the teaching of undergraduates on new graduates, for many of whom English is not a first language.

Quantity is another and a daunting aspect of the problem. The proportion of young people in the United States going on to higher education in science and engineering is only a half of what it is in Japan, while there is mounting evidence that demand is being constrained both by the high cost of higher education and the continuing poverty of high-school education at all but the excellent institutions. Can the United States continue to be astonished at the imbalance of its trade with Japan?

Who is to blame for this mess? The National Science Foundation. No other federal agency can be expected to do the job. Since the late 1960s, the proportion of the foundation's budget devoted to science and engineering education has declined from something close to 30 per cent to rather less than 5 per cent. The foundation finally gave up pretending to take seriously its statutory obligation to education at the undergraduate level in 1981, when its paymasters in the Office of Management and Budget (backed by the new White House) snuffed out all but one of its activities in that area. The one remaining programme provides a niggardly \$5 million a year for laboratory equipment in non-research institutions.

The result of this neglect is that university teaching has come to seem a chore to faculty members at many institutions, not an activity vital to the function of an institution. (There are some honourable exceptions, chiefly among the private universities.) Research, by contrast, brings its own rewards, both intellectual and (mainly in the United States) financial. What could the National Science Foundation do to help? Plenty. It could make awards, symbolic and monetary, for the preparation and dissemination of teaching materials, preferably making full use of the still largely untapped potential of computers. It could arrange (and pay for) national meetings and teaching workshops where good teachers would describe their techniques. It could make special grants to graduate students who display real talent for teaching. It could even expand its teaching equipment programme to include the research universities. Most important, it could renew its commitment to the notion that in an advanced democracy some comprehension of science, even by those who will not later become professional scientists, is a necessity, not a luxury.

A credible programme for science and engineering education at the undergraduate level would provide what is at present lacking, a subset of university teachers whose primary commitment would be to excellence in education. The amount of money needed is not large by US standards: \$50 million would be a useful start, increasing to perhaps twice that amount over a few years. Fortunately, all the indications are, the colossal deficit notwithstanding, that Congress is sympathetic to the idea and is willing to give the science foundation a shove in the right direction: it has demanded a plan of action by next March. Happily, the task force's study will be completed before then, in January, and not before time. By any analysis, the strength of the United States rests on its scientific and technical workforce, as do all of its hopes for the future. There is no known alternative to diligent study and excellent teachers. Americans should not need to be told that. □

Indian technology

Quest for self-reliance runs into trouble

New Delhi

SCIENTIFIC research in India is being steered by Prime Minister Rajiv Gandhi in a new direction intended to take the country into the twenty-first century — fast. Gandhi has made it clear that his government will not tolerate “rubbish science” or inefficient industry, and that all research must be productive. Indian scientists have been told not to waste time in reinventing technologies that can be purchased, but to work on those that others are not willing to sell. His adviser Mr L. K. Jha has gone a step further in suggesting that industries should not hesitate to import not only technologies but, if necessary, whole factories as well.

The plans for the overhaul of research and development and liberalized import policy stem from Gandhi's conviction that Indian scientists have been pampered for too long without being accountable and that industries have enjoyed protection for years at the cost of efficiency. Scientists are now being asked to deliver and industries to compete and catch up. “India had been pushing hard on science while neglecting the technology aspect”, Gandhi told a recent press conference. “Our research and development must be such that it really ends up with a finished product.” Gandhi feels that India's 750 million people are entitled to a better deal from Indian technologists. Since they have failed, there should be no objection to importing foreign technology.

The Council of Scientific and Industrial Research (CSIR) is the first agency to be asked to streamline its research. Other ministries that preside over research establishments will follow. Gandhi has directed CSIR to concentrate on a few key areas or “technology missions . . . instead of frittering away energy on a whole gamut of fields.” Following Gandhi's directive, CSIR director general Dr S. Varadarajan has undertaken the massive exercise of revamping the research activities of the council's 40 or so institutes. Several CSIR projects that lack definite goals are to be axed and pilot plants that have failed to generate interest in industry may be scrapped. About half of CSIR research is related to developing “import substitutes”. Now that import is freely allowed, such projects will face closure. Basic research will be a casualty. As far as CSIR is concerned, the only basic research permitted is that relating to “frontier areas” such as biotechnology or microelectronics.

Not everybody in CSIR is happy about the shape of things to come. “Our Prime Minister wants us to leapfrog to the

twenty-first century”, says a senior scientist. “But we cannot get there without mastering twentieth century technology.” Although a month has passed since the review began, few technology “missions” have been identified. “We can easily close down unproductive projects”, says Varadarajan, “but it is not easy to replace them with new ones.” CSIR has about 20,000 scientists and technicians, and it is an open question how to relocate those whose projects are to be axed. In the past, Indian industry has been more or less forced to adopt CSIR technology. According to one CSIR director, requests for processes from his laboratory dropped from 77 in 1984 to four in 1985. It is feared that with industries free to import the best from abroad there will be no more takers for CSIR knowhow.

The liberalized policy on the import of technology has also come under attack from sections of the scientific community inside and outside CSIR, as is shown by recent controversies over the move to import silicon and fibre optics technology. Critics believe the import policy is a negation of self-reliance, the hallmark of Indian science during the days of Jawaharlal Nehru and Indira Gandhi. “Indian research and development has been castrated”, says Baldev Singh, former chief of the technology utilization division of CSIR. “The country's laboratories appear almost like condemned institutions”, says Mr P. S. Deodhar of the Electronics Trade and Technology Development Corporation. There is a fear that industries based on indigenous technologies will be wiped out and that CSIR will face a slow death.

The fear is not unjustified. A cottage industry making silicon-carbide crucibles, and a factory that just started making “chlorosilane” using CSIR technology are facing closure as these products can now be freely imported. CSIR spent eight years developing SWAT-106, a chemical that improves oil flow through pipelines. But now four multinationals are planning to bring in their own technologies for a flow improver.

The *Hindustan Times* in a leading article said: “More and more public undertakings are hankering after foreign deals little bothering to find out whether appropriate indigenous knowhow exists. What is the idea of having big scientific establishments if their talents cannot be utilized locally?”

In the electronics sector, the clamour for foreign collaboration is at its peak. According to Deodhar, the import policy is luring businessmen “whose motive is commercial rather than technical”. Consumer electronics items with international

brand names are being marketed by Indian companies that have sprung up overnight. Some 29 giant computer companies from abroad are planning to enter the market, virtually sealing the fate of the state-owned Electronics Corporation of India Limited.

According to Prem Shanker Jha, a noted economist, the present import policy has spawned literally hundreds of “screwdriver units” in the electronics and automobile industries, “the two potentially biggest customers of capital goods industries”. There are three foreign collaborations for passenger cars, 10 for commercial vehicles and 14 for scooters and motor cycles. The government-owned Maruti car factory has rejected an Indian consortium's offer to set up an engine production line in favour of a Japanese company. Britain is setting up a plant near Bombay that will make 12,000 electric vehicles a year, preempting the electric vehicle project of the public sector Bharat Heavy Electricals Limited (BHEL), which is itself purchasing anticorrosion technology from a US company, ignoring processes developed by a CSIR institute.

“From one extreme of self-reliance, the pendulum has swung to the other”, says Baldev Singh, “and many of the current imports in the name of high-tech are really irrelevant to taking India into the twenty-first century.” Meanwhile, the number of foreign collaborations shot up from 440 in 1984 to 752 in the first six months of 1985 and foreign exchange reserves dipped from £4,800 million in April to £480 million by the end of July.

The peculiar situation in Indian science today is that heads of scientific agencies who privately worry about the technology import policy publicly remain silent or defend it. Many agree that the 1982 technology policy statement (TPS), which warned against indiscriminate import, has now been more or less torn to pieces. But Professor M. G. K. Menon, planning commission member and architect of TPS, will not comment on the new policy. And Rajiv Gandhi's government still swears by TPS. “The government has not deviated from the concept of economic self-reliance and the liberalized policy of import will not throttle indigenous research or industry”, the Prime Minister recently told newsmen. But sceptics, including the president of the Indian National Science Academy Dr C. N. R. Rao, who strongly opposed import of silicon technology, still believe that liberalized imports will stifle indigenous research.

Professor Abdul Rahman, a former CSIR director and science policy specialist, has warned that modernization cannot be achieved with imports of technology. “This”, he said, “has been amply demonstrated in Latin America which is saddled with debt, and in Iran where the Shah's hurry to modernize with massive technology inputs led only to political crisis.”

K. S. Jayaraman

Biotechnology

Diluted optimism at Osaka

Osaka

THE city of Osaka made its bid last week to establish itself as the Japanese "capital of biotechnology" with what was billed as "the first international conference on biotechnology". While that may be something of an exaggeration, the conference, thanks to generous support from the regional government, must have been about the first to hear experts from the Soviet Union describe their national biotechnology programmes alongside executives from many of the major biotechnology companies — Cetus, Calgene, Cytogen, Biogen, BioEurope and the like.

Within Japan, the Kansai region (which includes Osaka, Kyoto and Kobe) can fairly claim to have taken the lead in traditional rivalry with Kanto (the Tokyo-Yokohama megalopolis). Many of Japan's biotechnology companies have their research laboratories in the region, while Kyoto and Osaka University dominate basic research in molecular biology.

The state of international rivalry is harder to assess. Although Japan and the United States often regard themselves as the only players in the game, the European Communities representative was keen to scotch the view that Europe is not an active commercializer of new ideas. It was pointed out that most of the new (if still few) products of biotechnology were first produced in Europe: monoclonal pregnancy test kits (the Netherlands, 1980), biochemically produced human insulin (Denmark, 1982), recombinant amylase (Denmark, 1982) and recombinant α -interferon (Austria, 1985). Europe is also far ahead of the rest of the world in building new biotechnology research centres and providing integrated courses of study in the universities.

The UK outlook was more gloomy. Some praise was given to the establishment of new companies such as Celltech with "intellectual assets from the public sector and financial assets from the private". But there was a warning that lack of funds in Britain had pushed basic research to the brink of an irreversible decline; there may soon be nothing new to commercialize.

Japan's biotechnology industries are seen as building a better base for the long-term future. At present, commercial emphasis in the United States and Europe is on the production of new pharmaceuticals. That is where, in the relatively near future, completely new products, and new sources of profit, can be created. But it is in plant agriculture and food processing that the really huge returns on investment can be expected towards the end of the century. A very wide spread of companies in Japan has set up biotechnology research laboratories and knowledge of biotechnology is diffused through all kinds of in-

dustry. This diversity may pay dividends.

The Japanese tradition of focusing on production technology, rather than on an exotic end-product, is also at work in biotechnology. There is massive investment in exploiting biotechnology to improve existing processes — better bioreactors, enzymes, lipids and surfactants.

That biotechnology research is concentrated in large companies may also work in Japan's favour. A couple of years ago, Japan was bemoaning its lack of new venture capital while the United States was celebrating the host of innovative small companies into which the stock market was pouring money. But now there is a feeling that new venture businesses will not be big enough to capitalize properly

on their innovative achievements.

Not everyone was worried about international competition. The high level of cross-licensing agreements that has characterized the first phase of the growth of biotechnology, together with the diversity of uses to which biotechnology may be put, was seen by some as ruling out the kind of national rivalries found in electronics. But representatives from the South-East Asian countries were pessimistic. Dependent as many of them are on agricultural production, they stand to benefit most from biotechnology. But they have neither the facilities nor the manpower for large-scale research; recent developments, such as the successful patenting of new types of seeds, have them worried. The low attendance at lectures given by South-East Asian representatives must have helped confirm their fears.

Alun Anderson

Computerspeak

Hope for international standards

Washington

THE world is one year away from agreeing fundamental standards for controlling the communication of computers, according to the leader of the British delegation at a crucial set of meetings held during the past two weeks in North America. The negotiations involve more than half a dozen countries whose governments have been the most active since 1977 in attempting to arrive at a set of standards for inter-computer communications. The 300 or more delegates are here to discuss the latest development in Open Systems Interconnection (OSI).

The delegates are members of the sub-committees from the Geneva-based International Organisation for Standardisation and represent trade associations and government institutions. The main political impetus has come from the United States, Britain, Japan, West Germany, France, Canada and Scandinavia. The OSI meetings require the countries involved to move towards agreement on a comprehensive package of standards in seven distinct areas. These include the disciplines on how computers should be connected to telecommunications circuits, how those data can be stored and displayed, how one network can communicate with another to allow the "global addressing of computers", how errors in data in the network can be detected and corrected, how messages can be synchronized and their delivery controlled. The other standards enable types of data to be identified as they are transferred between computers over national boundaries and between different computer systems.

The sessions have been taking place in Toronto, Gaithersburg (Maryland), Syracuse (New York) and Durham (North Carolina).

Bryan Wood, leader of the UK delega-

tion, says "there is increasing support for a single standard independent of manufacturer". That indeed is the wish of the British government, which has refused to allow the UK telecommunications authority British Telecom to join with IBM to operate a data management network on the grounds that such a partnership would stifle competing systems and allow IBM's own standards to dominate. The IBM communications standard, recognized as a possible rival to OSI, is called System Network Architecture (SNA), and is still being developed in parallel with OSI.

The Europeans and Americans in particular hope that SNA will not hamper OSI development but that it will be used for intercommunications between IBM computers and that the OSI standard will be used for any other communication. The seven levels of the OSI standards, which are at various stages of ratification, are also meant to accommodate communications from any closed network to another.

Computers connected to each other within a factory or office complex would be examples of closed loops, called local area networks (LANs). Such networks will be able to be connected to each other using the OSI standards. They will be able to communicate with each other over an interconnecting network, possibly an international one, whatever the standard used at the local end. The Ethernet system originally devised by Rank Xerox and the recently announced IBM token ring, both forms of LAN, will be accommodated by the OSI standards, claim the delegates to the American conferences.

Bill Johnstone

Bill Johnstone, on sabbatical from his post as technology correspondent of The Times of London, will be writing on technology subjects for Nature during the next few weeks.

Eureka

Optimism in France . . .

"EXTREMELY satisfactory." That is how French research minister Hubert Curien described the results of the second ministerial conference on Eureka, the French-inspired initiative to develop new European high technology products, in Hannover last week.

The meeting decided on 10 projects, seven of which involve French participation. French officials were impressed that "the people who said no" to government finance of Eureka (which decoded means the British) are now saying "perhaps". Sir Geoffrey Howe, British Foreign Secretary, announced that British companies involved in Eureka projects could apply to the Department of Trade and Industry's £250 million annual "Support for Innovation Scheme", and said that he would not place an upper limit on the proportion of that fund that might be devoted to Eureka.

The French view, however, is in reality little different from the British. President François Mitterrand earlier this year announced that FF1,000 million would be devoted to Eureka in 1986, but it is now clear that none of this is new money, additional to the existing government research and development budget. According to Yves Sillard, the president of the French oceans and fisheries research council IFREMER and the man also responsible for the French Eureka projects, there will be no strictly new money in 1986. But more money could be provided later. As M. Hubert Curien, research minister, put it, there is now a line in the French research budget for Eureka, and this is politically defensible. M. Roland Dumas, foreign minister, stressed that the primary object of Eureka is to get new European technology on the market, thus emphasizing market pull rather than technology push — a point of view on Eureka adopted very early on by Britain.

According to Curien, the principal objective at Hannover was to make a list of projects, the first concrete plans, and that has now been achieved. With ten settled, fifteen more wait in the wings while detailed agreements are reached, and some 300 potential projects were on the table at Hannover, with many of them still in the running. "These are good days for Europe", Curien said.

These are, however, still honeymoon days for Eureka. No country or company has yet felt the pain of rejecting its own product in favour of one developed by outside companies under the Eureka label, which is what the logic of Eureka will ultimately demand. "There may be competing products even within Eureka", Curien added. "But we don't want the competition to lead to our losing the race with the United States or Japan."

According to Curien, Eureka will give

what would otherwise be merely a few bilateral or trilateral projects greater scale and greater speed than they would have outside Eureka. Eureka is accelerating what might have happened anyway, Curien said.

The controversial question of a secretariat for Eureka is to be decided, and the secretariat established, by 31 January next, but it should be small and "light", Curien said. Its location is not yet agreed, with countries divided roughly between those who support Brussels and a loose

Eureka

But who is going to pay?

Hannover

In the eyes of the West German host delegation, the Eureka conference was a limited success. Both foreign minister Hans-Dietrich Genscher and research minister Heinz Riesenhuber praised the second Eureka conference as a "starting point for a progressive Europe".

Nothing of real substance was actually decided, it is admitted, but the 36 foreign and research ministers at least left with good intentions. For example, all were agreed that the third Eureka conference should take place next year in London. Following traditional EEC practice, discussion of the most urgent problems has been put off until then. The ministers also decided in principle to create a "small and flexible Eureka secretariat" — Britain and West Germany eventually agreeing to this French proposal. West German Chancellor Helmut Kohl and France's President François Mitterrand are backing Strasbourg as a location for the secretariat but there is also support for Brussels.

All participants agreed that Eureka projects should only serve civil purposes — in stark contrast to the original plans of the French, formulated after their refusal to take part in the US Strategic Defense Initiative (SDI). West Germany rejected all military aspects knowing that only civil purposes of the planned research would find support in this country, although like Britain and Belgium, Germany intends to participate in both Eureka and SDI.

The ten projects most favoured in the initial list include all three suggested by West Germany. The ten are: high-power lasers for material processing; a computer network for European scientific research; investigation of atmospheric pollution independent of national borders; amorphous silicon for production of electricity from solar energy; superfast vector calculators; robots for the textile industry; biotechnology; optoelectronic production systems; diagnosis of sexual diseases, in particular AIDS; European standards for home and school computers.

connection with the European Commission, and those who do not. The French position is that the Commission's work and research should be treated as "complementary" to that of Eureka — which would argue, perhaps, for a Brussels base but not Commission domination of the Eureka secretariat.

Could multinationals be involved in Eureka projects? Yes, according to the French, if they had a base in Europe with independent power of decision. The role of countries beyond the 18 nations present at the Hannover meeting was not discussed, said Curien. Collaboration would not be ruled out but "Our object is clear. It is Europe, Europe." **Robert Walgate**

But there is still a cloud hanging over Eureka — there was no agreement about how to finance these ambitious projects. France made a gesture of offering to put FF 1,000 million toward Eureka projects; the West German government has not yet decided on the level of its support, preferring industrial companies to initiate and finance the projects.

Meanwhile criticism has come from some circles within German industry. If the industry is to finance the projects, why is it the politicians who meet to select them? The president of the German Trade and Industry Association, Wolf von Amerongen, went as far as saying that industry has absolutely no use for Eureka. The government disagrees. **Jürgen Neffe**

Refusniks appeal

LEONID Ozernoi, the Soviet Jewish astrophysicist who three years ago went on a protest fast coinciding with the eighteenth General Assembly of the International Astronomical Union (IAU) now seems to have lost faith in hunger strikes. In a message to the nineteenth General Assembly of IAU, which opens in New Delhi next week, Ozernoi says that his 1982 fast, and a similar, collective fast the following year with other refusniks seeking to emigrate to Israel had "disappeared without trace, like a black hole, in the depths of the Soviet emigration offices".

He adds that thanks to the support of friends in the Soviet Union and abroad, he has managed to stay out of prison or mental hospital and has not lost his job. But "nobody can be sure".

Ozernoi asks the participants at the IAU meeting to "raise [their] voices in protest against the retention of the refusnik scientists in the Soviet Union". To those who consider such appeals "inappropriate in the fight for peace", he says that silence will be interpreted as tacit support for Soviet attitudes towards the refusniks.

Vera Rich

Plant genes

Rifkin opens new campaign

Washington

JEREMY Rifkin, the campaigner against genetic engineering, has taken up a new cause: protection of plant germ plasm. In a suit filed in a federal court last week against the US Department of Agriculture (USDA), Rifkin, whose legal manoeuvres have held up field trials of genetically engineered microorganisms for two years, accuses USDA of "gross negligence" in looking after its collection of seed samples at Ft Collins, Colorado. The suit says the National Seed Storage Laboratory (NSSL) is so poorly managed and underfinanced that "thousands" of samples collected from all over the world are lost each year.

The Rifkin suit is the latest wrinkle in a highly politicized international dispute over the control of germ plasm. NSSL says it always makes available its samples (excluding narcotics) in response to a request from a qualified plant breeder, regardless of the request's country of origin. But Rifkin and his fellow plaintiffs, who include pressure groups from Malaysia, Spain and Canada, object to seed companies using the NSSL collection to develop usable breeds that may be highly profitable.

USDA drags feet

Washington

CONGRESSIONAL concern over the US Department of Agriculture (USDA)'s slow preparations to regulate biotechnology research is unlikely to be eased by a report published last week by the General Accounting Office (GAO). GAO found that in 1984 and 1985 together, the department will support biotechnology to the tune of \$40.5 million, including 87 projects that involve planned releases of genetically engineered organisms into the environment, most in the next five years. But GAO says that USDA was unable to provide information on the number, location and objectives of specific biotechnology research projects.

Denied official information, GAO conducted a questionnaire survey of institutions conducting USDA-supported research. The results of the researchers' own estimates of the risk of their planned environmental releases are, however, unlikely to surprise anyone. Asked whether their planned environmental releases would represent a reason for concern, 75 of the 87 contemplating such experiments checked the "no problem" box, nine the "very minor problem" box and three did not know. Asked what effort would be required to correct any problem that did arise, one solitary brave soul said a problem might be uncorrectable, noting, however, that this was the least likely outcome.

Tim Beardsley

The Rifkin version of this is that "industrialized nationals and transnational corporations attempt to seize control over the world's genetic resources in order to exert unchallenged power over the world's economy in the Genetic Age". The reason for the international involvement in a suit is that NSSL was for a time the principal repository of the seed collection maintained by the International Board of Plant Genetic Resources (IBPGR), an organization funded by the World Bank and the United Nations Food and Agriculture Organization (FAO) which has catalogued some 300,000 food crop varieties from around the world.

Rifkin says that NSSL, which is a backup safety repository, has haphazard collection policies; that most food crops other than a few staples are not collected at all; that samples are inadequately documented; and that genetic variation in the samples is lost because seeds are planted out too infrequently and in unsuitable locations. Rifkin wants USDA's total budget for germ plasm increased from its current level of about \$12 million to \$40 million, the figure recommended in an internal USDA study and in a study by the seed company Pioneer Hybrid.

While rejecting some of Rifkin's specific charges, USDA officials say they accept that the suit has been brought in good faith and readily agree that the situation at NSSL is far from satisfactory. Dr Charles Murphy, of USDA's Agricultural Research Service, admits that lack of storage space is becoming "a serious problem"; NSSL has put in a bid for extra funds towards a new storage building to replace one that will soon be full. And Murphy agrees that NSSL's lack of knowledge about many of the samples it has received is its "biggest concern"; efforts are now being made to implement a new computerized database management system. But the figures Rifkin quotes for the proportion of unusable samples make "no sense", according to Murphy. Many samples have not yet received "plant identification" documentation simply because NSSL has not yet been able to ensure the stocks are viable.

Dr Lewis Bass, curator of NSSL, also says that germ plasm storage should be a "higher priority programme", as it "will have far more impact on mankind than anything in the space programme". He says NSSL could use more money but that "we do a good job with what we get". Nevertheless, he agrees that the process called "enhancement" — getting the useful properties into strong varieties that can be used by breeders — has barely started.

IBPGR officials decline to comment on the specific charges being levelled at NSSL; they say that it is only by maintaining a strictly apolitical stance that IBPGR

has been able in the past to deliver samples to any country in the world.

The international aspect of the row is likely to come to a head this week at an FAO meeting in Rome. There, IBPGR's cherished apolitical stance may be seriously threatened: radical Third World countries have demanded that IBPGR be given a statutory foundation and guidelines that would make patenting of seed varieties illegal. In order to avoid becoming politicized, IBPGR may have no choice but to duck out of FAO's sphere of operations and become an independent scientific body.

Tim Beardsley

Bioengineering

Plan for field tests of bacteria

Washington

THE first environmental releases of genetically engineered organisms are expected to be approved this week by separate US government agencies. The Environmental Protection Agency (EPA) will give the nod to Advanced Genetic Sciences of Oakland, California, to conduct a field test of a bacterium modified to protect from frost damage, while the National Institutes of Health (NIH) will formally approve a proposal by Agracetus of Wisconsin to field-test a modified tobacco plant as a model for studying plant viral disease resistance.

Advanced Genetic Sciences had voluntarily sought approval for its experiments from NIH, but withdrew the application after legal complications bogged down a similar proposal from Stephen Lindow of the University of California at Berkeley.

The environmental impact assessment for the Agracetus proposal was signed by NIH director James Wyngaarden in September; NIH officials have in the meantime been deciding how much of the assessment will be made publicly available. The decision is that the entire document will be released, with only the details of the test location deleted. But it is unlikely that NIH will become involved in future environmental impact assessments of agricultural field trials; under new rules recently adopted, NIH can leave regulation to another agency if this is considered appropriate, and EPA is likely to deal with most such proposals in future.

Advanced Genetic Sciences has produced a modified strain of *Pseudomonas syringae* which has had deleted from its genome the gene for an ice-nucleator protein which in natural *Pseudomonas* is expressed in the cell membrane and triggers the growth of ice crystals when the temperature drops below freezing. Experiments with *P. fluorescens* will follow. Tests will involve treating strawberry plants with the modified bacteria.

Tim Beardsley

Israel

Medicine confronts Jewish law

Rehovot

TENSION between Israeli medicine and Jewish law, as embodied in the Torah, is growing. The immediate difficulty is that the practice of autopsies and transplants is being inhibited by the provisions of Jewish religious law that the body of a dead person must be treated with the utmost respect, must be buried quickly and in its entirety, and must "not be exploited".

Dr Mordechai Halperin, the physician and ordained rabbi who heads the Schlesinger Institute for Medical-Legal Research at the Orthodox-oriented Sha'arei Zedek Hospital in Jerusalem, says that one consequence is that autopsies cannot be carried out unless there is a "reasonable and immediate" prospect of the saving of human life.

During the past few years, the numbers of autopsies carried out at Israeli hospitals have declined sharply, chiefly as a result of the Anatomy and Pathology Law of 1981, a result of pressure from the religious parties, which requires that an autopsy requires not only the authorization of three

physicians (as of old) but also the approval of the next of kin, and which allows the procedure to be blocked by almost any relative, even a distant cousin. At Halperin's own hospital, autopsies are now few and far between, perhaps fewer than a dozen a year. But he says that there is no necessary conflict between medicine and Jewish law, and that autopsies are no longer necessary for the training of medical students or even for confirming physicians' diagnoses. At Sha'arei Zedek, where full-scale autopsies require the consent of the hospital rabbi as well as that of a senior physician, post-mortem biopsies are quite common.

Elsewhere, the proportion of hospital deaths followed by post-mortem examinations has fallen to 10 per cent, but there are even fewer in Orthodox Jerusalem. Professor Haim Lichtig, a pathologist at the Haifa Technion medical school, says that students did not participate in a single autopsy last term, and that textbooks, films and the examination of preserved organs are a poor substitute for the "the real thing".

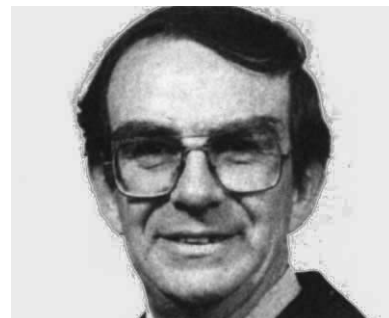
Pathologists such as Dr Nina Hurwitz of the Kaplan Hospital also argue that autopsies remain a valuable means of "quality control" in the practice of medicine, and that post-mortem biopsies or non-invasive diagnostic techniques such as CAT-scanning are inadequate alternatives; the discrepancy between clinical and pathological diagnosis has remained obdurately unchanged for 40 years, she says.

The availability of organs for transplantation is similarly constrained. Halperin says that Jewish law would treat as "murder" the removal of organs from a body whose heart was beating but whose EEG record was flat on the grounds that people had been known to recover from brain death. The removal of livers for transplantation would be permissible because artificial organs are not available, but kidney transplants are not always justifiable because kidney dialysis is possible. Halperin points out that life expectancy after kidney transplantation is not very different from that with dialysis.

The strict application of the law is not always accepted. One paediatrician who is religious considers autopsies necessary and has performed them himself. On kidney transplants, he suggests that one kidney should be removed from a dying person who may still be maintained by a respirator. Even some members of the Orthodox establishment appear to recognize that changes are required. Former Chief Rabbi Shlomo Goren, who also served as Chief Chaplain to the Israeli Army, is reported to have been furious that skin for transplanting to injured Israeli soldiers has had to be imported from overseas.

Nechemia Meyers

New RI director



PROFESSOR John M. Thomas, at present professor and head of the department of physical chemistry at the University of Cambridge, will become director, resident professor and director of the Davy Faraday Research Laboratory in the Royal Institution, London, on 1 October 1986, in succession to Sir George Porter, who has been nominated as president of the Royal Society. □

French agriculture

More help for farm lobby

THE French government has taken further steps to strengthen the research base of its largest industry, agriculture. The ministers of agriculture, M. Henri Nallet, and of research, M. Hubert Curien, have announced a new programme to double government assistance to private research projects. This follows the earlier announcement that the agricultural research council (INRA) will be reorganizing its research programme so as to adapt more quickly to the opportunities in biotechnology (see *Nature* 24 October, p.660).

The stimulus for the new programme, which will take government support for private research to FF110 million a year, was apparently the discovery that agricultural and food companies spend only 0.25 per cent of their turnover on research.

Interestingly, the new assistance programme will focus on the downstream part of agriculture. The four principal themes are nutrition and toxicology, product quality, fermentation and food engineering. INSERM is also to pay more attention to the epidemiology of nutrition, while there is to be a global strategy of research on food and health.

The emphasis on food safety seems to derive in part from the recognition that if the European Communities manage to establish a central scheme for food and drug regulation, along the lines of the US Food and Drug Administration, French food exports could be jeopardized. But the government seems also to be anticipating a weakening of European price support for food and other produce and the more competitive market that would then come into being.

Robert Walgate

Academies go west

Washington

THE US National Academy of Sciences is going west. Together with the National Academy of Engineering, it has been given \$20 million by the Arnold and Mabel Beckman Foundation to establish a 50,000 square foot west-coast building complex consisting of an auditorium and meeting rooms at Irvine, California. The Irvine Company has donated a seven-acre site adjacent to the Irvine campus of the University of California, worth \$6 million.

Arnold Beckman, now 85, is the chairman and founder of Beckman Instruments Inc., of Fullerton, California, and has recently made multimillion-dollar donations to both of his former colleges, California Institute of Technology and the University of Illinois (see *Nature* 17 October, p.568). The gift to the national academies includes an endowment to maintain and staff the building, which will be known as the Arnold and Mabel Beckman Center.

Beckman pointed out that California now leads the nation in memberships of both national academies. During fiscal year 1985, the academies and related institutions held more than 2,000 meetings in their east-coast buildings in Washington DC and Woods Hole, Massachusetts, but only 70 meetings were held west of Colorado. The new building will be used to support the full range of the academies' activities; Beckman has, however, expressed a particular desire that it should be used for the study of social and ethical issues arising from developments in science and engineering.

Tim Beardsley

UNESCO

Soviet move too late?

THE United Kingdom, which has given notice to withdraw from the United Nations Educational, Scientific and Cultural Organisation (UNESCO) by the end of the year, last week found an unusual ally in the Soviet Union, in the final days of UNESCO's Sofia conference. British delegates return home this week after a series of post-conference committee meetings and present ministers with recommendations for Britain's final decision.

UNESCO reaches its 40th birthday this month, but whether the battered organization will survive much longer has been called into question by the departure of the United States, along with 25 per cent of the organization's budget, at the end of last year. If Britain follows suit in December, as is likely, UNESCO will lose a further £5 million from its annual budget.

Timothy Raison, Minister for Overseas Development, spelled out Britain's worries at the beginning of the conference last month. The main concerns are that UNESCO is too top-heavy, bureaucratic, poorly managed and politicized, particularly in areas of disarmament policy (duplicating the work of other UN organizations) and government control of communications in the developing countries. Britain also considers that UNESCO should demand zero real growth in next year's budget and not increase contributions required of member states following US withdrawal. Specific proposals were presented to the conference as conditions for continued British membership.

Although several concessions were made to Britain in Sofia and the meeting was considered exceptionally harmonious by UNESCO standards, none of the major changes demanded by Britain were adopted. This always seemed unlikely; UNESCO's long-term strategy, the reason for convening the conference, was decided by its secretariat in Paris well before delegates arrived in Sofia — an example, in fact, of the centralized control that Britain is seeking to change.

A major stumbling block to the Western countries' demands for UNESCO's financial responsibility and an end to political patronage is acknowledged to be Amadou-Mahtar M'Bow, the director-general, a skilled manipulator, strongly supported by the developing countries. Hence last week's surprise development, in which the Soviet Union and its allies said that they would remove their support for M'Bow at the end of his term in 1987.

Whether this move will encourage the United States to open negotiations to re-join UNESCO remains in doubt. And the events at Sofia may well not be enough to cause a last-minute change of heart amongst British ministers. **Maxine Clarke**

Satellite broadcasting

Direct service standards tested

Washington

A RIGOROUS and novel series of tests of the technical attributes of three different broadcasting systems has just been completed at the Lewis Research Centre in Cleveland, Ohio, under the supervision of the National Aeronautics and Space Administration (NASA) by the US interests in satellite television. The tests will be crucial in assisting the United States to adopt soon a technical standard for its satellites.

The tests were conducted by communications specialists and untrained observers under the auspices of the Direct Broadcasting by Satellite Association (DBSA, which represents most of the parties involved in the new industry). DBSA will present a summary of its findings and recommendations by the end of the year to the Federal Communications Commission (FCC), which in turn will give its ruling on technical standards to the four operators who now have approval for DBS satellite construction. The ruling will be made next spring at the earliest.

The Lewis tests were comprehensive. They began in the middle of September and were concluded two weeks ago. The three systems under test were examined in four principal areas — video (picture), audio, teletext (information display) and scrambling (coding signal). Experts and untrained observers gave subjective views on the quality of television pictures resulting from each system.

The systems tested were two forms of NTSC, the method now used in the United States in terrestrial broadcasting, and a version of MAC (multiplex analogue component), the system favoured, in a different form, by the United Kingdom. The results will play a significant part in determining the technical composi-

tion of the US DBS satellites due for launch in late 1988 or early 1989.

Several companies have expressed an interest in operating DBS television services in the United States since FCC invited applications for slots in the geosynchronous orbit. In June 1983 a sectional meeting of the World Administrative Radio Conference (WARC), the body internationally responsible for assigning broadcasting frequencies and satellite positions, allocated eight satellite slots to the United States.

There are at present four approved DBS systems — Satellite Television (a subsidiary of Comsat), United States Satellite Broadcasting, Dominion Video Satellite and Hughes Communications Galaxy. A further six applications for DBS services are being considered. If all are approved, the technical system adopted by the DBS operators for transmission must be able to prevent satellites from interfering with one another. Some will be using the same frequencies, but will have been separated in different slots, while others will be clustered in the same orbit but operating on different frequencies, as each slot can carry 32 channels.

All four approved operators plan multichannel systems of six to 10 channels. One aspect of the tests is to find out the degree to which each signal format is susceptible to particular interference.

The scrambling tests are also of vital importance as the DBS services are likely to be supported by subscription and may need to be protected from piracy. The Lewis tests checked whether signals could be received and viewed by pirates despite the protecting code and whether there was any deterioration in the signal quality received by the legitimate viewer because of the code's presence. **Bill Johnstone**

Brenner to quit MRC post

BRITAIN'S top job in molecular biology will become vacant in 1987, when Dr Sydney Brenner plans to step down as director of the Medical Research Council's Laboratory of Molecular Biology at Cambridge. Brenner said last week that he plans to leave the laboratory at the age of 60, so as to be able to spend more time on his own research. He hopes to find himself a base elsewhere in Cambridge, and a small research group with which to pursue his present interests.

Brenner has been director of the MRC's best-known laboratory since 1979. An exile from South Africa by origin, Brenner joined the laboratory in 1957, when it was merely a barely tolerated collection of huts in a courtyard of the Cavendish Laboratory. During these early years, he worked with Francis Crick on the series of experiments that demonstrated that the genetic

code is a triplet code, and contributed substantially to the then rapidly deepening understanding of messenger RNA. More recently, Brenner has succeeded in opening up the study of the nematode *Caenorhabditis elegans*.

Brenner's spell as director of the Cambridge laboratory has been clouded by the consequences of a serious road accident in 1979, when a leg suffered a compound fracture. His colleagues marvel that his energy has been so little impaired.

The consequences of the vacancy at Cambridge will be considered by MRC at its meeting later in the month. The secretary, Sir James Gowans, said last week that all concerned are aware of the importance of the appointment, and that Brenner's successor will be sought on the international market. □

Climatic change with nuclear war

SIR—In August 1984, I pointed out in *Nature*¹ that the article by Turco and others² on nuclear winter did not arrive at a firm scientific conclusion, and is subject to numerous doubts. Fourteen months later, one of the authors, Carl Sagan³, moderated his statement, which had been reiterated in publications as diverse as *Parade*⁴ and *Foreign Affairs*⁵, that nuclear winter is a robust conclusion. His acknowledgement that the question must be further investigated is welcome. Unfortunately, the current article⁶ contains ambiguous and inaccurate statements.

As an example, Sagan³ quotes from the recent analysis by the International Council of Scientific Unions⁷. He fails to mention the conclusions of this study: that no nuclear winter is to be expected in the Southern Hemisphere, that considerable temperature declines could occur only in the northern latitudes in the summer and

the mid-continental areas. In mid-summer the effect will not be worse than those in fall or early winter. By contrast, the original TTAPS report, which Sagan calls a "valid first-order assessment", predicted a minimum hemispheric average temperature of -23°C .

To escape the danger of nuclear winter, Sagan suggested the reduction of stockpiles, which unfortunately requires faith in the reliability of Soviet compliance. Sagan has published in *Foreign Affairs*⁸

gave rise to soot, the darkening of the sky and to the consequent biological catastrophe.

The survival of approximately half the living genera witnesses the toughness of life on Earth. A reasonable estimate of both the energy released and the smoke

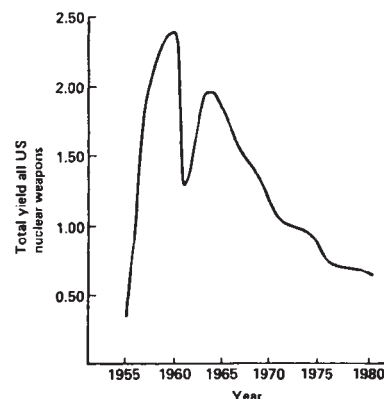


Fig. 3 The reduction in the total yield of US nuclear weapons over the years. Relative scale. 1972 = 1.00.

emitted following the impact of an asteroid surpasses the total energy and smoke released in a nuclear war by a factor of 1,000. In this scenario the massive amount of smoke may well have overcome the effects of atmospheric moisture.

By contrast, in a nuclear war, as was pointed out¹, the mass of smoke is 10^{-4} times the mass of H_2O in the atmosphere at the corresponding latitudes. Thus an "asteroid" winter seems plausible but hardly a nuclear winter.

Both Sagan and I are rightly worried and are rightly considering the damage a nuclear war may produce. We differ in many respects, but most particularly in my concern about the direct and the intended effects of nuclear war, including the local effects of prompt radiation, while Sagan has concentrated on side effects. These side effects are obviously dwarfed by an event for which there is geological evidence.

EDWARD TELLER

Lawrence Livermore
National Laboratory,
PO Box 808,
Livermore,
California 94550, USA

1. Teller, E. *Nature* **310**, 621–624 (1984).
2. Turco, R. P., Toon, O. B., Ackerman, T. P., Pollack, J. B. & Sagan, C. *Science* **222**, 1283–1292 (1983).
3. Sagan, C. *Nature* **317**, 485–488 (1985).
4. Sagan, C. *Parade*, 30 October, 4, 5, 7 (1983).
5. Sagan, C. *Foreign Affairs* **62**(2), 257–292 (Winter 1983/84).
6. Pittock, A. B. et al. *Environmental Consequences of Nuclear War, SCOPE 28 VI: Physical and Atmospheric Effects* (Wiley, Chichester, 1985).
7. News Release No. 424-83, 25 August (Office of Assistant Secretary of Defense (Public Affairs), Washington, DC 20301 1983).
8. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095 (1980); *Geol. Soc. Am., Spec. Pap.* **90**, 305 (1982).
9. Alvarez, W., Asaro, F., Michel, H. B. & Alvarez, L. W. *Science* **216**, 886 (1982).
10. Wolbach, W. S., Lewis, R. S. & Anders, E. *Science* **230**, 167–170 (1985).

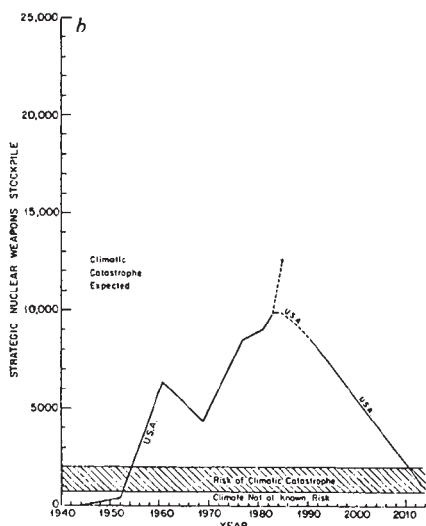
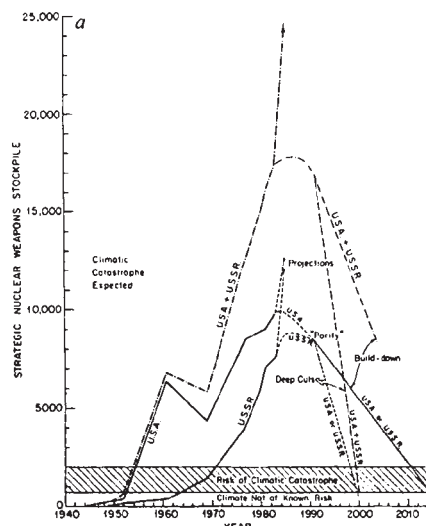


Fig. 1 a, Sagan's estimate of numbers of US and Soviet strategic nuclear weapons. b, An adaptation of Sagan's estimate of the numbers of US and Soviet strategic nuclear weapons.

data on the past and future numbers of nuclear weapons in the various stockpiles (Fig. 1). By contrast, the US government published⁹ relative figures on the time dependence of the American stockpile in regard to number of explosives and in regard to megatons (Figs 2, 3). It will be noted that the former have declined in the past two decades by 30 per cent, the latter by 75 per cent. This latter value is particularly significant in predicting worldwide effects. All of this was independent of disarmament negotiations.

To diminish the likelihood of the catastrophic impact of warfare, one should give most serious consideration to defensive measures. These may not only deter war but, in the horrible case of actual hostilities, the effect of nuclear war would be substantially reduced.

A basis of Sagan's claims is the work published by Alvarez and colleagues^{8,9} which adduces the presence of an iridium layer deposited 65 million years ago as evidence that an asteroid several miles in diameter hit the Earth. The subsequent dust excluded sunlight and led to the extinction of the saurians and many other species. On 11 October 1985, an additional paper was published¹⁰ where deposition of a terrestrial carbon layer was found to coincide with the iridium layer. The authors offer the plausible explanation that approximately 10^{31} erg was set free by a collision with an asteroid. This killed much of the vegetation, set forest fires and

CORRESPONDENCE

Origin of AIDS

SIR—In their recent review article in *Nature* (3 October, p.395), Wong-Staal and Gallo proposed that the HTLV-I virus originated in Africa and was brought to other endemic areas by commercial and slave trading dating back to the sixteenth century.

If this is true, I fail to see how the HTLV-I virus could have reached Alaska and Northern Scandinavia where it has been detected, and southern Japan where it is endemic. Japan was virtually cut off economically and politically from the rest of the world until the latter part of the eighteenth century. There was hardly any commercial trade between Japan and Africa before the twentieth century.

It is also important to emphasize that the presence of the HTLV-I virus in the vast expanse of continental Asia cannot be ruled out due to lack of epidemiological studies in the area.

Recent studies by the same authors now show that retroviruses highly similar to the human acquired immune deficiency syndrome (AIDS) HTLV-III virus occur not only in African green monkeys and macaques but also in some flocks of European sheep. These animals can all show AIDS-like symptoms with advanced infection. Hence the possibility of humans acquiring AIDS worldwide from animal sources cannot be overlooked. Indeed, such a possible origin of human AIDS has recently been proposed by Lewin².

JOSEPH ROSENIOR

University of Toronto,
Charles H. Best Institute,
112 College Street,
Toronto, Ontario,
Canada M5G 1L6

1. Gonda M.A. *et al. Science* **227**, 173–177 (1985).
2. Lewin, P.K. *Can. med. Ass. J.* **132**, 1110 (1985).

Biotechnology aids

SIR—In your report (*Nature* 6 June, p.448) on plans for the computing facilities at the European Molecular Biology Laboratory (EMBL), only passing reference was made to the "Biosequences" workshop held at Grottaferrata (near Rome) on 24–25 May.

EMBL's plans were one contribution in a lively and wide-ranging debate on criteria and guidelines for the selection, by the European Economic Community (EEC), of projects contributing to the development of user services and research infrastructure for biotechnology in Europe. The workshop, sponsored by EEC, the Italian National Research Council and the University of Bari and organized by the latter, was timely in relation to EEC's current funding programmes for bio-informatics and for the development of the market in high quality information services. It also offered interesting comparisons and contrast between the various national activities of the

United States, the United Kingdom, France, Japan, Italy and Ireland.

Of interest was the emphasis by both structural researchers (A. Lesk of MRC, Cambridge) and molecular evolutionists (C. Saccone of Bari, R. Grantham of Lyons) on the need for a comprehensive approach, integrating all the diverse sequence, structural and biological data available in constructing and refining models and hypotheses (Compare with Lesk, A. "Coordination of sequence data", *Nature* **314**, 318; 1985). The paper by B. Keil (Pasteur Institute) on the use of the proteolysis data bank to define the fixation sites of proteolytic enzymes illustrated the power of organized data.

These comprehensive and necessarily multidisciplinary approaches and the cost and difficulty of developing sophisticated and efficient software (for example C. Rawlings of ICRF on artificial intelligence approaches to structure prediction) and hardware (for example A. Coulson of Edinburgh on the application of distributed array processors to sequence comparison), lead to emphasis on better communication, networking and resource sharing between European centres. Training, users' clubs, electronic mail, software portability and on-line access were typical of the topics commended for support, complementing the centres and activities funded from national or EMBL resources. The workshop proceedings will shortly be available from the Research Directorate of the Commission of the European Communities (200 rue de la Loi, B-1049 Brussels).

C. FRONTALI

Istituto Superiore di Sanita,
Laboratory of Cell Biology,
Viale Regina Elena 299,
Rome, Italy

C. SACCONI

University of Bari,
Department of Biochemistry
and Molecular Biology,
Via Amendola 165/A
Bari, Italy

Creationism unloved

SIR—In his letter (*Nature* **316**, 184; 1985) C.K. Pallaghy, a senior lecturer in the school of biological sciences, La Trobe University, attributes the paucity of overt support for creationism by bona fide academics in reputable scientific departments to hostility to creationists by the hierarchy of their institutions and by their colleagues, and to fear of loss of respect and prestige generally. This attitude to creationists is hardly surprising in view of the methods they and their organizations adopt in promoting their fundamentalist religious views under the guise of "creation science".

Not content that our taxes are being used, through school grants, to support their fundamentalist Christian schools, they lobby vigorously for "creation science"

to be taught in the state schools and have already scored some success in Queensland. The creationists' acute aversion to evolutionism is such that they claim the *Genesis* creation stories to be the literal truth, despite their marked resemblance to Babylonian myths and to the mixtures of guesswork, poetic fantasy and folklore compiled by other races to account for the origin of man. So it is that Pallaghy, replying to a critic, had a letter last year in the *Melbourne Age* containing the passage, "first, does he realize that Noah's ark was immense (450×75×45 feet)? Its tonnage would have supported 4,000 fully grown African elephant bulls. There would have been plenty of room for all had God put young animals on board."

Everyone will agree with the creationists' right to believe whatever they like, but not with their wish to impose such beliefs on the young and vulnerable, in contradiction to the findings of the past several hundred years of intensely painstaking research — a priceless cultural heritage.

R.H. CLARKE

Meteorology Department,
Melbourne University,
Rosanna,
Victoria 3084,
Australia

PhD theses

SIR—I wholeheartedly agree with Beverly Halstead (*Nature* 29 August, p.760) concerning the relatively trivial return of a British PhD thesis volume from the comparatively substantial cost (invested?) for the purpose of its production.

The advocated prerequisite for the PhD award of publication of scientific articles arising from the thesis material is, in fact, similar to the system operated by many universities and colleges in both Scandinavia and the United States. Indeed, some overseas establishments demand only that the material be published in a journal without requiring a formal bound volume. No doubt some would object to the adoption of such systems in this country but at least the various funding bodies would be able to assess the value of their investment more readily as well as providing additional incentives to the PhD candidate.

Incidentally, I was enlightened by Beverly Halstead's mention of costs for technical assistance during the period of postgraduate research. This clearly does not apply in every case since neither myself nor the postgraduate students in this and other departments were offered the benefit of technical assistance. We were simply informed that the PhD was supposed to be "an individual effort".

I.C. KILPATRICK

University of Bristol,
Department of Pharmacology,
Medical School,
University Walk,
Bristol BS8 1TD, UK

Multiple sclerosis and viruses

Evidence that human T-cell lymphotropic viruses are associated with multiple sclerosis may lead to identification of viruses that cause the disease — or to an addition to the list of failed candidates.

It is not every day that the Faroe Islands are mentioned in the scientific literature, but no contemporary review on the causes of multiple sclerosis is complete without a mention of them. While the disease was unknown in the Faroes before 1940, a mini-epidemic was recorded in the years after British troops were garrisoned there in 1940. There seems little doubt that an infectious agent was involved; it is now conventional to suppose that multiple sclerosis can be initiated by viral infection in genetically susceptible individuals. But which virus or viruses are the culprits?

Frustration has marked the hunt. By some counts, twelve separate viruses have at one time or another been implicated as initiators of multiple sclerosis. But in no case has the case been made to stick. So the first evidence of the frequent exposure of multiple sclerosis patients to a new virus type is bound to excite an equal mixture of interest and scepticism. Publication of the data, moreover, is bound to quicken the pace of the further research that will show whether the new type of virus is genuinely an initiator of multiple sclerosis, or whether it will join its dozen or so predecessors on a list of suspects. It is in this spirit that we publish, on page 154, the tantalizing but inconclusive evidence from Hilary Koprowski, Robert Gallo and their colleagues of an association between multiple sclerosis and the human T-cell lymphotropic viruses (HTLVs).

This provisional group of viruses comprises three members. HTLV-I was the first to be discovered, in 1980, and is the cause of adult T-cell leukaemia discovered in Japan. HTLV-II is a variant of HTLV-I; and HTLV-III (also known as lymphadenopathy-associated virus, LAV) is the cause of acquired immune deficiency syndrome (AIDS). Although HTLV-III lives up to this name in some ways, notably a propensity to infect human T lymphocytes, its nucleic acid sequence is not related to that of HTLV-I or HTLV-II. Another view, now strengthening, is that HTLV-III is more related to the lentiviruses than to the other HTLVs. Evidence comes from both structural comparisons and the fact that the HTLV-III can be found in brain tissues of certain AIDS patients (Shaw, G.M. *et al.* *Science* **227**, 117; 1985). Lentiviruses are animal viruses that cause progressive demyelinating diseases of the brain and spinal cord. Indeed, the disease caused in sheep by visna virus, the most studied of the lentiviruses, is often

said to be the best naturally occurring model for multiple sclerosis.

Inevitably, therefore, attention has been turned to the possibility that HTLV-III or a related virus is the cause of multiple sclerosis. Ironically, the evidence of Koprowski *et al.* is if anything more suggestive of an association between multiple sclerosis and HTLV-I or HTLV-II than with HTLV-III. Indeed there is no uniform response among the patients studied to tests for any of the three HTLV types; the authors say this conceivably indicates the presence of one or more entirely new HTLV-like viruses, possibly including a virus that combines some of the features of all three known HTLVs.

Leaving aside the identity of the viruses responsible for the signals detected in the tests, what and how strong are the signals, and to what extent can they be interpreted as evidence that the viruses in any way cause the disease? The signals are of two distinct types: the presence in the blood and cerebrospinal fluid of antibodies that cross-react with HTLV proteins, and the presence in cells cultured from cerebrospinal fluid of HTLV RNA. The former indicates that the viruses have been present, the latter that they are still present and active.

There are three problems in interpreting the antibody data. First, although it is clearly shown that the concentration of antibodies that cross-react with some HTLV proteins is greater in most serum and cerebrospinal fluid samples from multiple sclerosis patients than in various controls, such antibodies are not invariably present. Second, no consistent picture emerges as to which HTLV the antibodies are most closely related, although HTLV-I is much more frequently and usually more strongly implicated than HTLV-III. Third, and most important, it is already known that elevated concentrations of antibodies to measles and Epstein-Barr virus are frequently found in the cerebrospinal fluid of multiple sclerosis patients, throwing doubt on the significance of raised antibodies to other viruses.

For the data that indicate the presence of RNA that is related to HTLV RNA in cultured cerebrospinal fluid cells, at least there is no consistency of HTLV type. In the 4 out of 8 cases where RNA has been detected, it is related, though not very closely, to RNA of HTLV-I but not HTLV-III. When present, the RNA is detected in less than one cell in ten thousand,

which may at first seem too low to be of significance, but this rate is similar to the frequency of HTLV-III RNA in T cells of AIDS patients. Moreover, it is a general observation in chronic virus infections, including visna and measles, that only 0.1 to 1.0 per cent of cells in any tissue contain viral genetic information. Much more of a problem is that the controls are so far limited to two healthy subjects, neither of whom had HTLV-related RNA in the cells of their cerebrospinal fluid.

When Ashley Haase and his colleagues first reported measles virus RNA in several multiple sclerosis brains (*Science* **212**, 672; 1981), the data, together with those on elevated antibodies to measles virus in multiple sclerosis, made the measles virus a strong candidate as the initiator of the disease. But further investigations revealed that measles virus RNA can be detected almost as frequently in the brains of patients with other neurological and non-neurological diseases as in those with multiple sclerosis, weakening the case for the involvement of measles virus in multiple sclerosis, though not excluding it.

Clearly, more data, and particularly more controls, will be needed before the case for an HTLV-type virus as the initiator of multiple sclerosis can be realistically assessed: the National Multiple Sclerosis Society in New York has already organized several laboratories to begin participating in studies of the putative association with HTLV. There will also be intense efforts to try and isolate the virus or viruses that are the source of the HTLV-I-like RNA. Until that can be achieved, there can be little progress in understanding the variability of the HTLV-related antibodies in multiple sclerosis.

Even if a virus can be proved to be implicated in multiple sclerosis, it is likely to be merely the initiator of a chain of immunological events leading to demyelination of nerve fibres rather than a direct cause. It seems probable that the initial viral infection somehow triggers an autoimmune reaction against components of the myelin sheath of nerves. "The baroque complexities of contemporary immunology in its application to multiple sclerosis", as Byron Waksman puts it in summarizing current research on page 104, will require considerable unravelling before it is clear how the disease progresses from the triggering event.

Peter Newmark

Astronomy

When is a star a superstar?

from C. Martin Gaskell

How massive and luminous can a star be? The traditional upper limit to the mass of a stable main-sequence star is about 60 times that of the Sun, but superstars with masses thousands of times greater than this have been seriously considered for the power sources in quasars (or at least as possible progenitors of the supermassive black holes currently favoured as these power sources). The object R136a in the Large Magellanic Cloud has been claimed to be such a superstar. The latest observations¹, however, show that it is not a single object but a small compact cluster of bright stars, some of which nonetheless have a luminosity that exceeds that of the brightest known star in our Galaxy.

The dominant feature of the Large Magellanic Cloud — an irregular companion galaxy to our own Milky Way — is a gigantic region of star formation, visible to the naked eye, and known as 30 Doradus (the name comes from the 1712 star catalogue of John Flamsteed, the first astronomer royal). This nebula — often called the Tarantula — is the biggest region of star formation in our local group of galaxies. It is over 3,000 light years across and contains half a million solar masses of gas. In its centre are some extremely bright stars, dominant among which is HD38268. More commonly known as R136 (it was number 136 in a list of stars published by the Radcliffe observatory² in 1960), it is putting out enough ultraviolet radiation to ionize almost the entire Tarantula nebula.

What is R136? Studies at the beginning of the century showed it to have the spectrum of a very hot O star. In 1973, N. R. Walborn suggested that it was not one star but a small compact cluster of very luminous stars³. He argued that compact systems of a handful of bright stars within a light year or so were often found in the very centres of giant star-forming regions. The four stars known to amateur astronomers as the Trapezium, in the Orion nebula, are the nearest example. Walborn argued that if a system like the Trapezium were transported to the distance of the Large Magellanic Cloud, it would resemble a single star on photographs. Further work showed that R136 does have several components⁴ but two groups suggested that the brightest component (R136a) was a single massive superstar with a mass of 2,000–4,000 solar masses^{5,6}. Such a star would be very interesting as it would far exceed the conventional upper limit to the masses of normal stars. This limit, first derived by A. S. Eddington⁷, and named after him, occurs because the intensity of radiation becomes so great that radiation pressure exceeds gravity. R136a is in fact

losing mass at an enormous rate: Cassinelli, Mathis and Savage estimate it loses the equivalent of the entire mass of the Sun in 3,000 years⁸.

Better spatial resolution was needed to see if R136a is indeed a single superstar. The resolution of photographs taken by telescopes from the Earth's surface is set by turbulence in the atmosphere, and the very best photographs of R136a, taken in Chile, have a resolution of 0.7 arcseconds⁹. It was C. E. Worley, of the US Naval Observatory, who realized that R136a had already been resolved by visual micrometer measurements made 60 years ago¹⁰ but overlooked by modern spectroscopists. Worley, himself, made further visual observations in 1983 and published the results in a one page paper¹¹ (perhaps the last time anyone will publish a visual micrometer observation of a double star in the *Astrophysical Journal Letters*). Worley's observations confirmed that R136a is actually two stars separated by 0.49 arcseconds with one (R136a₁) being twice as bright as the other (R136a₂). R136a₁, if a single star, would still be a very impressive object of perhaps 750 solar masses⁷. But is it a single star?

The issue has finally been settled by G. Weigelt and G. Baier using the technique of holographic speckle interferometry¹.

From 8,000 short-exposure pictures of R136, taken with the Danish 1.5 m telescope at the European Southern Observatory, they were able to reconstruct the true image of R136a because of the presence of the point sources R136b and R136c in the same field (holographic reference stars). Their conclusion is that R136a₁ is actually a double star with a 0.10 arcsecond separation, and that there are at least 8 stars within a circle 1 arcsecond in diameter. Weigelt and Baier have therefore shown that R136 is indeed a complex of stars, much like the Trapezium in Orion, as Walborn had suggested.

Although nobody now believes 30 Doradus contains a 2,000 solar mass superstar, R136a remains highly interesting. Its three brightest components are each probably of more than a hundred solar masses, brighter than any star in our Galaxy. Many questions about the birth, life and death of such stars remain. □

1. Weigelt, G. & Baier, G. *Astr. Astrophys.* **150**, L18 (1985).
2. Feast, M. W., Thackeray, A. D. & Wesselink, A. J. *Mon. Not. R. ast. Soc.* **121**, 337 (1960).
3. Walborn, N. R. *Astrophys. J. Lett.* **182**, L21 (1973).
4. Feitzinger, J. V., Schlosser, W., Schmidt-Kaler, Th. & Winker, C. *Astron. Astrophys.* **84**, 50 (1980).
5. Cassinelli, J. P., Mathis, J. S. & Savage, B. D. *Science* **212**, 1497 (1981).
6. Eddington, A. S. *The Internal Constitution of the Stars* (Cambridge University Press, 1926).
7. Chu, Y.-H., Cassinelli, J. P. & Wolfine, M. G. *Astrophys. J.* **283**, 560 (1984).
8. Innes, R. T. A. *Southern Double Star Catalogue* (Johannesburg Union Observatory, 1927).
9. van den Bos, W. H. *Ann. Leiden Observatory* **14** Pt 4, 1 (1928).
10. Worley, C. E. *Astrophys. J. Lett.* **278**, L109 (1984).

C. M. Gaskell is in the Astronomy Department, University of Texas, Austin, Texas 78712, USA.

Embryology

Interacting systems in amphibia

from Hugh Woodland and Elizabeth Jones

AMPHIBIAN embryos have long been a favourite experimental material for embryologists because their large size and independent development makes them particularly amenable to techniques like micro-injection and grafting. They are therefore particularly suitable for finding out how the variety of cells in the very early embryo differentiate. This kind of information is virtually non-existent for *Drosophila*, another developmental system much in favour at the moment. *Drosophila* work relies mainly on mutants and those available affect higher orders of pattern, rather than the basic events of cyto-differentiation. Thus, although the power of genetics cannot be applied to the amphibians, they still have a very important place in embryological research. Moreover, new life has very recently been injected into the study of amphibian embryology by the introduction of cloned DNA and antibody reagents which distinguish particular types of differentiation.

Until now, it has been necessary to rely

on purely morphological criteria for identifying tissues in amphibian embryos and, in many kinds of experiment, tissues may be very abnormal and their identification hard to justify, especially to those without extensive histological experience. Sometimes, as where cell division is blocked with agents that disrupt the cytoskeleton, the correct morphology of the tissue could not possibly appear. Cloned DNA molecules or antibodies (mostly prepared from *Xenopus* embryos) are being introduced to circumvent these problems. The DNA gives very accurate quantitative information but, since nobody has yet developed a viable method of *in situ* hybridization for early stages of amphibian development, spatial information is limited. The reverse is true of antibodies, which may readily be used for the immunohistological staining of sections.

How are the main tissues of the embryo generated? At the time of nervous system formation, all vertebrates have the structure shown in Fig. 1c. The considerable

tissue diversity is generally believed to originate from heterogeneity in the egg's cytoplasm, which is subsequently partitioned into different cells during cell division (cleavage). The small number of initial cell types thus formed are thought to interact with each other to generate further complexity. The original example of such an interaction, the induction of the nervous system, was discovered by Spemann and Mangold over 50 years ago. Induction is the process in which the pathway of development of tissue A is changed by interaction with tissue B while that of B is unaffected. In the case of the nervous system, ectoderm is the responding tissue; in isolation it forms only epidermis, but placed with dorsal mesoderm (the future notochord and segmental voluntary muscle blocks) it forms nervous system (Fig. 1*g-i*). This invites the question of what causes the mesoderm to develop, and whence comes its dorsal-ventral polarity. This problem, which has received considerable attention lately, is especially interesting because the mesoderm is responsible for so much of the internal complexity of the body.

It was originally thought that the mesoderm arises from a special type of cytoplasm in the zygote, but there is no unequivocal evidence that the very early embryo contains more than two cell types, the ectoderm and the endoderm. The former develops from the pigmented half of the egg, the latter from the non-pigmented yolky half. In the 1960s, the field was revolutionized by Ōgi¹ in Japan and Nieuwkoop² in Holland. Nieuwkoop, who has not received the recognition appropriate to the importance of the work, concluded that the mesoderm is mainly or wholly formed by an induction in which the yolky endoderm induces its formation from ectoderm (Fig. 1*a*). This was shown by experiments in which animal and vegetal cells, neither of which form mesoderm in isolation, were removed from an embryo and grafted together (Fig. 1*d, e*). The result was the generation of mesodermal tissue (Fig. 1*f*), from which it can be concluded that a variety of mesodermal tissues may be induced from ectoderm that would normally form epidermis and nervous system. This basic result was soon supported by Nakamura's group in Japan³.

Very recently, cell-specific markers have provided a sophisticated confirmation of this process. John Gurdon, Tim Mohun and their colleagues in Cambridge have used cloned actin genes and an antibody to pick up the muscle component of mesoderm formed during mesodermal induction⁴; in Jonathan Slack's lab in Mill Hill, London, antibodies which react with muscle, epidermis and notochord have been used for the same purpose⁵.

These experiments have also more accurately established the dynamics of mesoderm induction. For example, even

in those experimental situations in which tissues have to re-establish communication, induction can take as little as 1.5–2.5 hours⁴. Furthermore, responsiveness of the ectoderm to induction and the ability of the endoderm to induce is lost in the early gastrula. It is important to know these kinds of details for at least three reasons. First, losses of responsiveness happen to be the first detectable signs that cells have adopted a particular path of development. Second, knowing when responsiveness and inductiveness occur should help in identifying the molecules responsible. Third, charting loss and gain of these properties will help in understanding the organ-building processes of early development. During the extensive morphogenetic movements of gastrulation, cells completely change their

that the agent is naturally relevant to mesoderm formation, even in the chick. The same is true of putative neural inducers isolated from amphibians⁶. In both cases, the next step is possibly to purify the putative inducers and generate monoclonal antibodies against them. These could then be tested for their ability to block induction and used in a strategy to clone the genes.

Oddly enough, the precise time when natural mesodermal induction occurs is still unknown. It might seem simple to take embryos at different stages, remove the cells normally expected to form mesoderm, and to find the earliest stage at which they develop, into, say, muscle. Unfortunately, throughout the likely period of induction, it is not possible to recognize any of the cells concerned. Fate mapping indicates that the general region

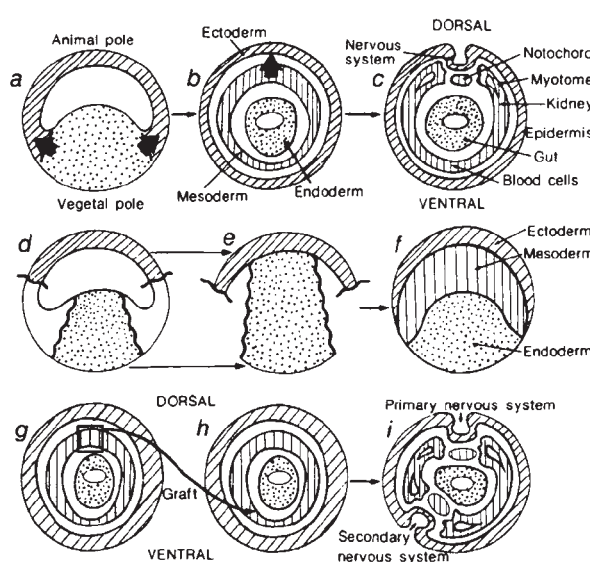


Fig. 1 Diagrammatic transverse sections of amphibian embryos in normal development (*a-c*) and in grafting experiments (*d-i*). Blastula stage embryo (*a*) with precursor ectoderm and endoderm, which interact (arrows) to form mesoderm between them. *b*, After gastrulation: dorsal mesoderm induces (arrow) overlying ectoderm to form nervous system. *c*, Neurula stage: the main body organs are mapped. If the animal cells and vegetal cells are removed (*d*) and combined (*e*), they interact to generate mesodermal tissues (*f*). If dorsal mesoderm is removed from one embryo (*g*) and placed ventrally in another (*h*), the latter develops two sets of axial structures (*i*).

neighbours. If cells maintained a fixed response to a certain set of stimuli, or if the stimuli were always present, abnormalities could result. New phases of mesodermal induction can, and almost certainly do, occur later, for example during formation of the tail⁶ and perhaps the muscular cells of the iris.

We need to know if there is any unity in these processes, but there is unlikely to be much advance on this front until the molecular basis of at least one type of mesoderm induction is established. Progress has been slow in this direction. Although usable assays exist, they indicate that embryos at the appropriate stage contain only minute amounts of the substances that cause mesodermal induction. However, agents that cause induction have been isolated from more abundant sources, such as late-stage chick embryos or guinea pig bone marrow. A few micrograms of the chick agent have recently been purified and 10 nanograms of this 13,000 M_r protein have been shown to be sufficient to induce a nervous system⁷, though it remains to be proved

which forms mesoderm also contains cells which normally form endoderm or ectoderm, and which are capable of interacting to generate mesoderm up till at least the gastrula stage. Thus it is not possible to prepare an appropriate tissue fragment that is demonstrably incapable of undergoing new induction, even if the normal interaction has already occurred.

This same problem bedevils attempts to show that part of the mesoderm might be produced by an area of egg cytoplasm containing localized mesodermal 'determinants'⁸. Although the overall region that forms mesoderm in isolation contains probable ectoderm and endoderm precursors¹⁰, improved cell-marking techniques show that most of the mesoderm comes from the pigmented half of the embryo¹¹, which forms only epidermis in isolation. This strongly supports induction as the natural mechanism of mesoderm formation, since the pigmented region must develop into mesoderm only through the proximity of the yolky half of the embryo.

The mesoderm itself is obviously non-

homogeneous; its dorsal side contains the notochord and somites, as well as the neural-inducing ability, while the blood islands are ventrally located (Fig. 1b, c). Nieuwkoop, himself, first clearly showed that this vital dorsal-ventral polarity is imposed by the endoderm², experiments which have been elegantly refined by Bob Gimlich and John Gerhart in Berkeley¹². In its turn, the dorso-ventral asymmetry in the endoderm arises from cytoplasmic movements in the egg, polarized by the entry of the sperm — as discovered in the 1930s but now explored in some depth by the Berkeley group¹². The result is that the dorsal endoderm is formed on the side opposite sperm entry.

The simplest explanation of subsequent events is that there are three kinds of cell interaction that bring about mesoderm induction. First, (though not necessarily temporally so) there is mesoderm induction, which alone produces solely ventral mesoderm. The ventral nature of this effect is shown by obliterating the polarizing effect of the sperm on an embryo, for example by UV irradiation, whereupon a radially symmetrical embryo develops; most of the mesoderm forms blood cells, and none of it forms notochord or somites¹³. Second, the dorsal endoderm somehow instructs the overlying future mesoderm itself, to become dorsal, that is notochord and somite. This can be shown by transplanting pairs of blastomeres at the 32-cell stage¹², or by a modification of the operation shown in Fig. 1d-f, in which endoderm from future dorsal and ventral sides was transplanted separately^{14,15}. The third kind of interaction occurs within the mesoderm itself. Dorsal mesoderm influences the rest, perhaps through generating a morphogen gradient, to produce the dorso-ventral pattern of structures outlined in Fig. 1c. This is

demonstrated by the famous Spemann and Mangold experiment in which the dorsal mesoderm is grafted into a ventral position (Fig. 1g, h); a second set of notochord, somites and nervous system is generated, and the result is a 'siamese-twin' embryo (Fig. 1i). The influence of the graft on its surroundings is shown by the fact that the secondary set contains host cells as well as those of the graft¹⁵.

Sorting out the detailed cell and molecular biology of these interactions, as well as the generation of the initial cell heterogeneity by cytoplasmic determinants, needs simple experiments (like the grafts shown in Fig. 1d-f) that can be adapted to provide routine assays for the molecules involved. Fortunately, the development of such assays will be greatly facilitated by the cell-specific markers currently being developed, not least because they will make it possible to bypass time consuming histology. □

1. Ōgi, K.I. *Sci. Rep. Tohoku Univ. Ser. IV (Biol.)* 33, 239 (1967).
2. Nieuwkoop, P.D. *Curr. Top. Dev. Biol.* 11, 115 (1977).
3. Nakamura, O. In *Organiser* (eds. Nakamura, O. & Toivonen, S.) 179 (Elsevier, 1978).
4. Gurdon, J.B., Fairman, S., Mohun, T.J. & Brennan, S. *Cell* 41, 913 (1985).
5. Dale, L., Smith, J.C. & Slack, J.M.W. *J. Embryol. exp. Morph.* (in the press).
6. Spofford, W.R. *J. exp. Zool.* 107, 123 (1948).
7. Tiedemann, H. in *Biochemistry of Differentiation and Morphogenesis 23 Colloquium Ges. Biol. Chem.* (ed. Jaenicke, L.) 275 (Springer, Berlin, 1982).
8. Janeczek, J. *et al. Eur. J. Biochem.* 140, 257 (1984).
9. Gurdon, J.G., Mohun, T.J., Fairman, S. & Brennan, S. *Proc. natn. Acad. Sci. U.S.A.* 82, 139 (1985).
10. Heasman, J., Wylie, C.C., Hausen, P. & Smith, J.C. *Cell* 37, 185 (1984).
11. Cooke, J. & Webber, J.A. *J. Embryol. exp. Morph.* (in the press).
12. Gerhart, J.C. *et al. Proc. R. Soc. Lond.* 307, 319 (1984).
13. Malacinski, G.M. in *Pattern Formation* (eds Malacinski, G.M. & Bryant, S.V.) 435 (1984).
14. Botterbrood, E.C. & Nieuwkoop, P.D. *Wilhelm Roux Arch.* 173, 319 (1973).
15. Slack, J.M.W. *From Egg to Embryo* (Cambridge University Press, 1983).

Hugh Woodland and Elizabeth Jones are in the Department of Biological Sciences Warwick University, Coventry CV4 7AL, UK.

Human disease

Mechanisms in multiple sclerosis

from Byron Waksman

THE splendor of the great stairs and ceilings of the Würzburg Residenz, created by Balthazar Neumann and Tiepolo, provided a fitting backdrop for a recent discussion* of the baroque complexities of contemporary immunology in its application to multiple sclerosis (MS), a crippling disease of young adults.

Susceptibility to MS seems to be under multigenic control, with links to *HLA-D* locus genes *DW2 (DR2)* on chromosome 6 (C. Jersild) and possibly to IgG allotype genes (*Gm-1, 17:21*) on chromosome 14 (J. Pandey *et al.*). A role for genes governing the T-cell receptor subunits and/or vasoactive amine sensitivity, suggested by

the results of animal studies (see below), remains conjectural, as is the significance of increased frequency of spontaneous and induced chromosomal translocation in MS cells (G.R. Sutherland).

The evidence that MS in susceptible individuals is initiated by viral infection, at least in some cases, is increasingly persuasive. For the epidemic in the Faeroe Islands that followed the arrival of British troops in 1940, it seems that approximately 2 years of exposure to the putative infectious agent was necessary and that the period from infection to onset of clinical disease was about 6 years (J. Kurtzke). Susceptibility was low before puberty. In informal discussion, it was mentioned that MS elsewhere is associated with infections

(measles, mumps, rubella, Epstein-Barr virus) caught significantly later in age than in HLA-matched controls (A. Compston, O. Andersen). In a well-studied cohort of patients, over 25 per cent of exacerbations were associated with such viral infections; from two weeks before to five weeks after infection, the exacerbation rate was increased more than threefold (W. Sibley).

Viruses such as measles, rubella, and varicella have been shown to sensitize T lymphocytes to myelin basic protein (MBP) in some subjects; this sensitization is very frequently associated with 'post-infectious' encephalomyelitis (R.T. Johnson). In a study of such encephalomyelitis associated with rubella, multiple T-cell clones obtained from cerebrospinal fluid showed equal reactivity with the virus and with MBP (P. Marquardt). Three independent computer searches have indicated the possible chemical basis for this cross-sensitization by demonstrating the presence of homologous peptide sequences in encephalitogenic regions of MBP and in antigens of measles, influenza, canine distemper, Epstein-Barr virus and several papova and adenoviruses (R. Fujinami, U. Jahnke *et al.*, G. Stoner).

A key point of current research is to determine the exact location where specifically-immune T Lymphocytes, circulating in the blood, first recognize myelin (or other) antigen, 'presented' in conjunction with major histocompatibility antigen (Ia). In agreement with the findings in animal studies, Ia has now been demonstrated on the endothelium of blood vessels within MS lesions and on activated astrocytes near the edges of lesions, as well as on the population of infiltrating macrophages (U. Traugott). MBP has also been demonstrated at these locations. Thus presentation sufficient to trigger T cells in the blood, or such T cells as enter the perivascular space, may occur at the luminal surface of the endothelium or the foot processes of the astrocytes. In older lesions, presentation may take place on the cell bodies of astrocytes, or other glia, and on the surface of macrophages.

The early MS lesions, triggered in the white matter of the central nervous system by these reacting T cells, show vascular injury with leakage of fluid, invasion of the tissue by a mixed population of lymphocytes and macrophages, and myelin destruction (C. Raine, H.L. Weiner). T lymphocytes, mostly of the T8⁺ class but with some T4⁺, predominate. Haematogenous macrophages play the principal role in myelin destruction at the advancing edge of the lesion (J.W. Prineas) by a mechanism that seems to involve receptor-mediated endocytosis. That the ligand for the receptor may be antibody is suggested by immunoglobulin 'capping' on many of the actively phagocytic macrophages. But extracellular lysis of myelin is also seen: oligodendrocytes at the edge of the lesion look healthy, while those in contact with lymphocytes or macrophages show mor-

Multiple Sclerosis — Current Experimental Work and Some Clinical Perspectives Würzburg, 29–31 August, 1985.

phological evidence of cell damage.

T cells found in the cerebrospinal fluid may reflect events in the central nervous system. Both T8⁺ and T4⁺ types are present and many carry markers characteristic of long-term cell lines (Weiner). There is one report of cells specific for MBP in very acute cases of MS (R. Lisak and B. Zweiman), but most studies of specificity suggest that the T cells are polyclonal and that many are MHC-autoreactive (A. Salmi, G. Birnbaum). The cerebrospinal fluid also contains raised levels of γ -interferon, even during clinical remission (R. Hirsch). The B cells and plasma cells of the lesions and the cerebrospinal fluid show a polyclonality that is comparable to that of the T cells and affects all the major immunoglobulin classes (K.P. Johnson, H. Link). The much debated 'disappearance' of T8⁺ cells from the peripheral blood early in acute attacks (J. Antel and B.G.W. Arnason) has been attributed either to their actual migration into the target tissue (Weiner) or to modulation of their phenotypic markers by, for example, prostaglandins released from activated monocytes in the circulation (P. Dore-Duffy, J. Merrill). A striking new finding is the absence, in most patients, of cytotoxic T-cell precursors specific for measles virus (H. McFarland). There is no comparable deficiency for influenza virus.

For many of these crucial new findings, the study of animal models with conditions that resemble MS has led the way. Experimental allergic, or autoimmune, encephalomyelitis (EAE), induced by immunization with MBP or proteolipid protein, may take a chronic relapsing or progressive form, particularly if immunization is at the time of weaning (equivalent to puberty) (Raine, D. McFarlin, F. Lublin, T. Tabira). Certain temperature-sensitive mutants of the JHM virus produce a subacute or chronic inflammatory demyelinating disease in rats that is associated with sensitization against MBP; lymphocytes from these animals produce typical EAE in normal recipients (H. Wege, R. Watanabe). Theiler's murine encephalitis virus (TMEV) produces similar disease in SJL mice and certain other strains (H. Lipton); although there is an immune response against myelin, the continuing reaction may be directed at persistent viral antigen (Lipton, P.L. Lampert).

The use of recombinant inbred mouse strains has shown there to be multigenic control of susceptibility to the MS-like disease, as well as to its intensity and character, in both EAE and TMEV models. For EAE, both high-responder and low-responder strains have been found (Lublin). The high responders, most frequently associated with *H-2^s*, readily develop chronic relapsing or progressive disease. Earlier studies in acute EAE had mapped the responsible *H-2* gene to *I-A* (R. Fritz). Genes controlling vascular sensitivity to vasoactive amines are also important (S. Linthicum). For the TMEV sys-

Hypothetical sequence of pathogenetic events in multiple sclerosis

Acute phase

Systemic virus infection (measles, varicella, influenza)

Primary (cross-reactive) immunization to MBP

Circulating T lymphocytes specific for MBP

Systemic interferon (α , β , γ) response

Activation of vascular endothelium (and astrocytic foot process?) by viral particles and/or by γ -interferon

Expression of Ia in vascular endothelium body-wide and in CNS

Acute phase events in CNS vessels

Expression of Ia and presentation of MBP

Dual recognition by MBP-specific T cells

Activation of T cells leading to nonspecific secondary inflammation

Leakage of fluid producing oedema

Invasion of tissue by monocytes (macrophages), T cells and B cells

CNS tissue damage by oedema and invading cells

Demyelination by activated macrophages

Damage enhanced by local antibody forma-

tion, complement from circulation, free oxygen radicals (?)

Arrest of conduction, due to fluid pressure, loss of myelin, and perhaps T-cell lymphokines and free MBP

Chronic phase

Activated T cells produce IL-2 and γ -interferon locally

Activation of astrocytes by γ -interferon

Expression of Ia, release of IL-1 and prostaglandins by astrocytes invading macrophages

Proliferation, synthesis of GFAP, gliosis

Activation of Ia-reactive T cells in response to Ia, IL-1 and IL-2 in lesions

Persistent foci of Ia-reactive T cells

Continuing nonspecific secondary inflammation

Systemic virus infection-reactivation of process

Astrocyte activation by systemic γ -interferon

Renewed activity of local Ia-reactive cells

Enlargement of existing lesions

New lesion formation, as in acute phase

MBP, myelin basic protein or other myelin antigens; Ia, major histocompatibility complex antigen (Class II); IL-1 and IL-2, interleukins-1 and 2; GFAP, glial fibrillary acidic protein; CNS, central nervous system.

tem, an important genetic control, related to the intensity of the T-cell mediated response against viral antigen, maps to *H-2D*; another seems to be related to the gene encoding the β -chain of the T-cell receptor.

Lines or clones of T cells specific for MBP produce acute or chronic EAE when activated (with mitogen or specific antigen) *in vitro* and transferred to normal syngeneic recipients (I.R. Cohen, H. Wekerle). The delay of 72–96 hours before the first inflammatory cells appear in the central nervous system (R. Meyermann) can be accounted for by the time required for the transferred cells to produce heparanase, which permits them to penetrate the vessel wall and enter the tissue (Cohen), and to produce γ -interferon, which activates vascular endothelial cells and astrocytes. The delay includes time for both cell types to synthesize and express Ia, which can be demonstrated on macrophages infiltrating lesions (W. Fierz, Sobel).

An important new finding is that purified endothelium from brain vessels can be stimulated *in vitro* by activated T-cell supernatants or purified γ -interferon to express Ia and present antigens such as MBP to specifically sensitized T-cells (McFarlin). Similar data for purified astrocytes have been reported (B. Fontana). The importance of presenting myelin antigen in association with Ia is clearly established in experiments with cell lines transferred from, or to, either of two rat strains and their F1 hybrid (C. Linington). JHM virus particles, however, are capable of directly activating astrocytes without mediation by γ -interferon (P. Massa). Manoeuvres which stimulate systemic γ -interferon and/or reduce suppressor T-cell activity trigger relapse, even in normally resistant animals, such as Lewis

rats that have recovered from acute EAE. An important confirmatory finding is that MBP-specific suppressor cells are present in the spleens of sensitized rats during periods of remission but not during exacerbation (W. Lyman and C. Brosnan).

A hypothetical sequence of events in MS, which merges findings reported at the symposium with reports in the recent literature, is presented in the table. Several points of attack for potential therapies are suggested by the scheme. It would be worth testing the value of monoclonal antibodies directed at phenotypic markers or characteristic activation antigens of appropriate T-cell subsets; at the T-cell receptor or its subunits, particularly idiotypes related to specific myelin antigens; or at Ia (Weiner). Hybrid or toxin-linked antibodies may be advantageous. Other leads come from studies of EAE. Vaccination with MBP-specific lymphocytes, inactivated with X rays or mitomycin C, induces resistance (idiotypic tolerance?) to actively induced EAE (Wekerle); lymphocytes inactivated by hydrostatic pressure (which changes the physical state of membrane glycoprotein) are particularly effective, inducing a resistance that is transferable with T cells (Cohen); prolonged treatment with MBP plus galactocerebroside not only turns off chronic EAE but promotes remyelination, perhaps due to oligodendrocyte stimulation by antibody to galactocerebroside (Rodriguez).

All in all, participants at the meeting were imbued with a sense of optimism that both the complex scientific problems related to MS and the practical issues of its prevention might be nearing solutions. □

Byron H. Waksman is Director of Research Programs, National Multiple Sclerosis Society, 205 East 42nd Street, New York 10017, USA.

Cell metabolism

Organization in the cell soup

from Hans V. Westerhoff

METABOLITES moving directly from one enzyme to the next, enzymes sticking to each other and to membranes, intracellular water diffusing freely, energetics and kinetics not matching actual metabolism: in a recent workshop on *The Organization of Cell Metabolism**, these and related observations fell into a coherent picture. Rather than a thick, homogeneous soup consisting of metabolites and proteins at high concentration, the cell sap emerged as a bouillon with vermicelli. The vermicelli corresponds to a lattice of structural proteins to which cellular enzymes are adsorbed; metabolites move directly between the enzymes adsorbed next to each other, and where available, membranes can take over the role of the lattice.

As examples of the organization of metabolism through the dynamics of enzyme kinetics itself, glycolytic oscillations in yeast (B. Hess, Max Planck Institute, Dortmund) and oscillations in cyclicAMP concentration and movement in *Dicyostelium discoideum* (A. Goldbeter, Université Libre Bruxelles) were discussed. The models assumed the enzymes to be dissolved in a homogeneous solution of metabolites. However, the highlight of the workshop was the understanding of the spatio-temporal organization of metabolism arising from the breakdown of this very assumption. For instance, the *in vitro* desorption of NADH from α -glycerolphosphate dehydrogenase was reported to be greatly accelerated by excess lactate dehydrogenase, suggesting direct transfer to that enzyme. In general, direct transfer of NADH seems to occur between enzymes that bind it in the *syn* conformation and those that bind it in the *anti* conformation, which makes sense mechanistically (S. Bernhard, University of Oregon). In some cell types, most of the hexokinase is bound to a pore-forming structure in the mitochondrial outer membrane; glycerokinase added to a mitochondrial preparation does not effectively compete with the bound hexokinase for mitochondrial ATP (J. Wilson, Michigan State University).

These are extreme cases of 'metabolic channelling' — the transfer of metabolites from an enzyme molecule producing them to an enzyme molecule consuming them without equilibration with a 'pool' of the same metabolite. Such channelling can be either 'static' (where the two enzymes form a steady complex), or 'dynamic' (where metabolite transfer is contingent on the enzymes colliding).

In vitro, association of enzymes is often rather weak and dependent on low ionic strength (L. Kanarek, Vrije Universiteit

Brussel). But if the enzymes also bind to a support, the ensuing cooperative binding may suffice for physiological conditions. The inhibitory effect of oxaloacetate on the binding of citrate synthase to submitochondrial particles, suggests that association of enzymes may well be a point of metabolic control, which would explain why such association tends to be weak (P.A. Srere, Veterans Administration Medical Center, Dallas).

In view of this lability of enzyme supercomplexes, assay methods for channelling in intact systems have to be either delicate, such as the 'gentle' opening of cells that is necessary if cellular enzymes are subsequently to be shown to co-elute on chromatography columns (P.D.J. Weitzman, University of Bath), or based on functional properties implied by channelling. Several examples of the latter were presented. First, there is more rapid increase in specific activity of intracellular proteins than in that of intracellular amino acids upon extracellular addition of labelled amino acids (D.N. Wheatley, University of Aberdeen). Second, lactate is produced in toluene-permeabilized cells from glucose-6-phosphate but not any other added glycolytic intermediate (D.B. Kell, University College of Wales). Third, intermediary metabolites injected into frog oocytes fail to lead to the isotope dilution expected if metabolism were not channelled (T. Ureta, Universidad de Chile). Spatially inhomogeneous NAD(P)H fluorescence, after similar injections into mammalian cells, has been observed (E. Kohen, University of Miami). Another approach focused on the one-to-one relationship expected in static

channelling between enzymes producing a metabolite and enzymes consuming it, and on the implication that an inhibitor of the former enzymes would not affect the titre (with respect to the pathway flux) of an inhibitor of the latter. In membrane-linked free-energy transduction, this type of experiment provides an indication of channelling (D.B. Kell).

Nuclear magnetic resonance experiments have suggested that the inside of a cell resembles a thick but homogeneous soup. It is not obvious why channelling should occur in such a soup. Now, however, an alternative interpretation seems to prevail. The apparent seven-fold reduction in the self-diffusion coefficient of water is now thought to reflect the presence of barriers to (otherwise free) diffusion. Thus, shrinking of cells causes a reduction of translational, but not rotational, motion of a water-soluble electron-spin resonance probe (A.M. Mastro, Pennsylvania State University) and quasi-elastic neutron scattering suggests that, when measured over a few Angstroms, the diffusion coefficient of intracellular water is reduced only by a factor of three rather than seven (J. S. Clegg, University of Miami; C. Hazlewood, Baylor College of Medicine, Houston).

Against the background of structural information on the cytomatrix, there emerges a picture of an ultrafine structural lattice ramifying throughout the cytoplasm. Most of the intracellular protein is loosely attached to the lattice and water is free to move the short distances in between its threads. As was shown for artificially co-immobilized enzymes (A. Friboulet, Université de Compiègne), the proximity of the adsorbed enzymes will result in channelling. □

Hans V. Westerhoff is in the Laboratory of Molecular Biology, the National Institute of Arthritis, Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

Plate tectonics

Palaeozoic longitude problem

from A.G. Smith

GREAT strides have been made in quantifying global geology for Mesozoic and Cenozoic time — the past 250 million years. This success has depended on magnetic surveys of the ocean floor and on measurements of the magnetic field preserved in rocks on the continents. But the situation is quite different for Palaeozoic time (245–570 Myr), because virtually all the Palaeozoic ocean floor has gone down subduction zones, to be recycled into the mantle, thus destroying half of the critical data needed to reposition one continent relative to another. But a recent paper by L.P. Zonenshain, M.I. Kuzmin and M.V. Kononov (*Earth planet. Sci. Lett.*

74, 103; 1985), of the P.P. Shirsov Institute of Oceanology in Moscow, presents some new Palaeozoic reconstructions which are considered by the authors to have brought a solution within sight.

By appealing to plate tectonic theory and using Euler's theorem — which states that the motion of a tectonic plate can be represented by rotation about an axis through the Earth's centre and a point on the surface — geologists can make precise reconstructions of the relative positions of the Mesozoic and Cenozoic continents from the oceanfloor data. By assuming that the Earth's magnetic field has always been a dipole whose axis lies parallel to

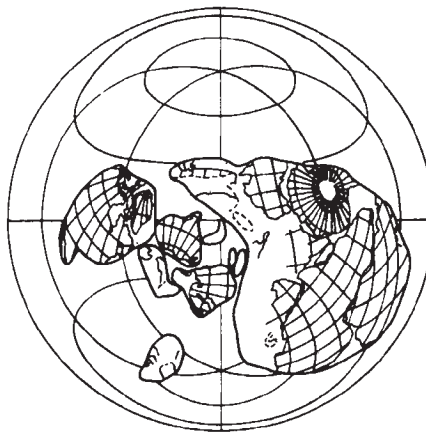
* 31 August – 4 September, 1985, Hanstholm, Denmark.

the axis of rotation, they can put these reconstructions into their former latitudes from the inclination and declination of the magnetic field preserved in continental rocks of the appropriate age. There is no difficulty at all in estimating the relative longitude separation of the major continents for the past 180 Myr or so, as the longitude inherent in palaeomagnetic measurements is resolved by the ocean-floor data.

When two continents are fixed with respect to each other, their relative positions can be found by superimposing their apparent polar wander (APW) paths, which in theory should be identical. Gondwanaland has been reconstructed by this method, but then it was independently known that Gondwanaland was a single continent during the time interval for which the APWs were compared. If the motion between the continents is small, then applying the method will give a reconstruction that will contain relatively small errors. The problem in using this method for an arbitrary period in the past is knowing how much of the difference between two similar APWs is due to errors of measurement and how much to relative continental motion.

By adopting a criterion of minimum motion between the continents in time, Mesozoic and Palaeozoic reconstructions have been made on the basis of palaeomagnetism alone, and are not dissimilar to those based on ocean-floor spreading (Irving, *E. Nature* **270**, 304; 1977). More recently, R.A. Gordon, A. Cox and S.O'Hare (*Tectonics* **3**, 499; 1984) argued that Euler poles can be inferred from long APWs of the continents, though this use of the palaeomagnetic data contains hidden assumptions that should be more closely examined.

An alternative method of making reconstructions is to use the hotspot reference frame as a means of repositioning the continents. Provided hotspot traces are available for all the continents, they can be placed in the same frame independently of any other data and the longitude problem can be solved. R.A. Livermore, F.J. Vine and I (Geophys. J. R. astr. Soc. **79**, 939; 1984) have tested this idea as applied to the past 90 Myr and have shown that the differences in latitude inferred from the palaeomagnetic reference frame and the hotspot reference frame are less than 7°, with the pole enclosed by, or within a degree or two of, the α -95 circles of confidence for that interval. For the previous 90 Myr the difference progressively increases to nearly 20°, possibly because the hotspot frame for the Jurassic and Cretaceous periods is not well known. In cases of incomplete data, hotspots could be used to orient one or more continents, the remainder being oriented using palaeomagnetic Euler poles. If there are more than enough data, these methods can be combined appropriately to find the best solution. This is the approach the



A reconstruction of the positions of the continents in Cambrian time (roughly 550 Myr ago). Modified from Zonenshain, L.P. *et al.* (*Earth planet. Sci. Lett.* **74**, 103; 1985).

Russian group have attempted to use.

The problem with using the hotspot frame for Palaeozoic reconstructions is that no Palaeozoic hotspot traces had previously been suggested. What Zonenshain and his co-workers now propose is that three examples of igneous activity within Eurasia represent continental hotspot traces for the time periods 400–365 Myr, 400–320 Myr and 280–130 Myr. Taken together with approximate Palaeozoic reconstructions from palaeomagnetic Euler poles, these traces would enable the relative motions of the Palaeozoic continents to be found in the hotspot frame. The most striking difference between previous reconstructions and theirs is in the Cambrian. The continents are not only much closer together but are also in a different order, with Siberia lying between western

Europe and North America (see the figure).

One of the major problems in this kind of work is the scarcity of reliable and well-distributed palaeomagnetic data. Granted that, one question is what is the significance of the palaeomagnetic Euler poles? In essence, the method assumes that any long APW consists of arcuate segments joined together by cusps (Gordon *et al. op.cit.*). The cusps are interpreted as representing important periods of plate reorganization, the arcuate segments as periods of steady motion. The existence of cusps, or at least of sharp changes in the trend of an APW, is well known and almost certainly reflects a discontinuity in the motion of the continent concerned, such as its separation from, or collision with, another continent. But what of the arcs between? Obviously, any arcuate path on a sphere can be described geometrically by a pole and an angle. What is the evidence that such poles refer sufficiently closely to the same reference frame, so that they are useful guides for reconstructions of past continents as opposed to reference frames that differ significantly from continent to continent?

A second obvious question is whether there are any hotspot traces on other Palaeozoic continents. If not, does the systematic age progression in the igneous activity in Eurasia have a different origin, such as one related to progressive rifting? Whatever the eventual answers to these questions, the paper by Zonenshain *et al.* will stimulate much further research and discussion. □

A.G. Smith is in the Department of Earth Sciences, University of Cambridge, Cambridge, CB2 3EQ, UK.

Palaeoanthropology

The earliest *Sivapithecus*?

from Eric Delson

THE search for a fossil ancestor (or even an extinct close relative) of the human lineage has met with mixed success over the past few decades. During the 1960s, the view that *Ramapithecus* was an ancestor of humans as contrasted with the great apes became dogma in anthropology¹, questioned by few workers and those usually on invalid grounds. Sharper attacks on this position characterized the 1970s, as details of molar enamel thickness, dental proportions and wear patterns revealed little difference between this taxon and *Sivapithecus*, its contemporary in 9–13 million year old rocks in India and Pakistan^{2,3}. Finds of more complete specimens attributable to *Sivapithecus* confirmed the similarity of these taxa and moreover permitted the inference that this expanded *Sivapithecus* group shared derived facial and dental features with *Pongo*, the living orangutan^{4,5}. This

left the fossil ancestry of human and African apes, joined phylogenetically by numerous generic and biomolecular studies⁶, open to question. Minority alternative views have arisen which suggest that *Sivapithecus* is ancestral (or at least closely related) either to all great apes and humans⁷ or to orangutans and humans alone⁸.

New evidence on the morphology and temporo-spatial distribution of *Sivapithecus* is therefore of great interest in this discussion. On page 173 of this issue, Leakey and Walker⁹ report and briefly describe new fossils from Buluk, Kenya, which they assign to an indeterminate species of *Sivapithecus*. This information is potentially exciting because the fossils are found in sediments that are more than 17 million years (Myr) old, as documented on page 175 by McDougall and Watkins¹⁰, far older than the earliest Eurasian fossils assigned to *Sivapithecus*. Based on this

age, Leakey and Walker now accept that *Sivapithecus* could well be the common ancestor of all later large apes, but realize that it tells little about the divergence of either the African lineage (Homininae: chimp, gorilla, human) or Asian (Ponginae: orangutan) lineage implied by molecular studies. But how accurate is their "unequivocal" identification and subsequent phyletic assessment, and what more (or less) can be inferred from these fossils?

Before these questions can be approached, it is useful to review the Buluk primates and other known occurrences of Miocene large apes. The fossil primate material from Buluk is rather scrappy, but four species are said to be present: three 'apes' of varying size and a more common species of Old World monkey. Although not cited by Leakey and Walker, the monkey has recently been described in detail by Meave Leakey¹¹, who assigned it to *Prohylobates*, a genus otherwise known only from slightly younger deposits in Egypt and Libya¹², rather than to *Victoriapithecus* of the Kenyan? Early to Middle Miocene (or about 17–15 Myr)^{13,14}. This apparent affinity with northern rather than eastern African taxa is mirrored in the remainder of the Buluk faunal list, which suggests either that the environment was more open than other Early Miocene sites nearby (perhaps akin to that of the Kenyan Middle Miocene sites Maboko and Fort Ternan) or that there was an influx of immigrant species.

Of the Buluk hominoid primates, a few specimens are referred to a small species, most to the new *Sivapithecus*, and an isolated tooth and dentulous mandible fragment to *Kenyapithecus wickeri*. The last taxon, intermediate in size, is best known from Fort Ternan but has also been reported from Maboko and nearby Majiwa¹⁵ and from Emuruiem (Nachola)¹⁶. There is still some controversy over the name and number of species of *Kenyapithecus* to be recognized but, despite previous synonymization with *Ramapithecus*, most authors now accept that *Kenyapithecus* is a distinct African mid-Miocene genus. While Pickford¹⁵ has grouped *Kenyapithecus* with the *Sivapithecus*–*Pongo* clade, Andrews¹⁷ places it on a separate lineage predating the Ponginae–Homininae divergence, because of its lack of derived *Pongo*-like features.

A number of other fossil 'populations' may also be relevant to understanding the Buluk specimens. In Asia, *Sivapithecus* is well documented in the Indo-Pakistan Siwaliks at about 12.5 Myr from remains with the distinctive facial characteristics seen in more complete younger specimens but, on dental evidence¹⁸, may extend back to over 14 Myr; recently, remains as young as 5.5 Myr have been assigned to this genus¹⁹. In Pakistan, the most complete fossils date to about 8 Myr. A large sample (probably representing a single species with high sexual dimorphism, contra Wu and Oxnard²⁰), of similar age from

China has also been placed in *Sivapithecus* and/or *Ramapithecus*²¹ but may require reassessment. The wide interorbital spacing and distinctive high-crowned incisors are not comparable to Siwalik *Sivapithecus*, although dental morphology is similar.

Sivapithecus has also been identified in Turkey and eastern Europe as old as 15 Myr, mainly on the basis of thick molar enamel, which Martin has shown to be a more complex character than previously thought²². The identification of *Sivapithecus* in Greece (see ref. 23) has been more widely accepted²⁴. De Bonis and Melentis²¹ have now reported that the maxillary morphology of the Greek fossils, which they term *Ouranopithecus*, is more similar to that of hominines than *Sivapithecus* (see ref. 5). Kelley and Pilbeam²⁵ have suggested that the morphology thus revealed was derived and they have linked *Ouranopithecus* to hominine ancestry, but I prefer the de Bonis and Melentis interpretation that hominines share the ancestral morphology of this region, with the pongines like *Sivapithecus* being derived. Given that *Ouranopithecus* does share other derived features with *Sivapithecus*²⁴, it seems best to retain all the Eurasian thick-enamelled Miocene apes in a single clade.

In the light of this diverse assemblage of African and Eurasian Miocene larger hominoids, it is somewhat surprising that Leakey and Walker⁹ only compare the Buluk fossils to *Proconsul* and the Siwalik *Sivapithecus*. The novel faunal elements would suggest a wider search for comparative samples. Based upon the description and a brief examination of casts (courtesy of Walker), it seems to me that at least one alternative interpretation of the Buluk material is possible, if not more likely — that it represents a large *Kenyapithecus*. Pickford¹⁵ has presented new metrical data on *Kenyapithecus* which support the idea that the Buluk *Sivapithecus* is the male of a species whose female is identified there as *Kenyapithecus*. For example, his estimate of the length of a male upper canine of *Kenyapithecus* is slightly greater than that given for the Buluk male, while molar lengths are comparable to, or slightly smaller than, those estimated for Buluk from roots; Leakey and Walker⁹ do not give the direct measurements of the few relatively complete molars they have.

More to the point is the comparison of maxillofacial and mandibular morphology and enamel ultrastructure. The Buluk mandible is stated to be "remarkably similar" to an edentulous one from the Pakistani Siwaliks, but while this is so by comparison with *Proconsul*, there is little to differentiate among the Eurasian forms, all of which share superior and inferior transverse mandibular tori as well as a deep and thick corpus. The Buluk maxilla is not specifically compared with any Siwalik specimen, but the described morphology does read like that seen in several Asian *Sivapithecus* faces. It is not clear,

however, that the described pattern is unique to *Sivapithecus* or is as well developed in Buluk as in the Eurasian fossils. An enamel thickness of 1.5 mm is given for the cuspal region of one broken, but little worn, upper molar, somewhat less than the 2–2.5 mm recorded for Pakistani *Sivapithecus*²⁶. Martin²² has shown that ultrastructural detail (unknown in this case) as well as size normalization, is needed to make an accurate estimate, of the enamel pattern, but the Buluk fossils may present the "intermediate-thin" pattern, predicted by Martin but not yet observed. (Neither *Kenyapithecus* nor *Dryopithecus* has yet been analysed.)

The essence of Leakey and Walker's comparative analysis is that since the Buluk fossils are not *Proconsul*, they must be *Sivapithecus*. While the first part of this syllogism is unquestioned, their conclusion is by no means inescapable. The lack of even superficial comparisons with *Kenyapithecus*, with the Moroto maxilla (which has long been assigned to a large species of *Proconsul*²⁷, but may be of Middle Miocene age⁵ and thus rather younger than most of that taxon) or with *Dryopithecus* is surprising. It would seem, in fact, that an allocation to *Kenyapithecus* would be both more biologically parsimonious and reasonably supported by the morphological evidence. That *Kenyapithecus* is positioned close to the divergence of Ponginae and Hominae is thus reaffirmed, but the implications for the broader phylogeny of large hominoids in terms of palaeontology or molecular-clock calibrations are uncertain. □

1. Simons, E.L. & Pilbeam, D.R. *Folia primatol.* 3, 81 (1965).
2. Simons, E.L. & Pilbeam, D.R., in *The Functional and Evolutionary Biology of Primates* (ed. Tuttle, R.L.) 36 (Aldine, Chicago, 1972).
3. Greenfield, L. *Am. J. phys. Anthropol.* 50, 527 (1979).
4. Andrews, P. & Tsekaya, I. *Palaeontology* 23, 85 (1980).
5. Ward, S.C. & Pilbeam, D.R. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R.L. & Corucini, R.S.) 211 (Plenum, New York, 1983).
6. Andrews, P. *Nature News and Views* 314, 498 (1985).
7. Kay, R.F. & Simons, E.L. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R.L. & Corucini, R.S.) 577 (Plenum, New York, 1983).
8. Schwartz, J.H., *Nature* 308, 501 (1984).
9. Leakey, R.E.F. & Walker, A. *Nature* 318, 173 (1985).
10. McDougall, I. & Watkins, R.T. *Nature* 318, 175 (1985).
11. Leakey, M.G., *Folia primatol.* 44, 1 (1985).
12. Delson, E. *Geobios* 12, 725 (1979).
13. von Koenigswald, G.H.R. *Foss. Vertebrates Afr.* 1, 39 (1969).
14. Szalay, F.S. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
15. Pickford, M. *J. hum. Evol.* 14, 113 (1984).
16. Ishida, H., Pickford, M., Nakaya, N. & Nakaya, Y., *Afr. Study Monogr.*, Kyoto Univ. Suppl. 2, 73 (1984).
17. Andrews, P. in *Ancestors: The Hard Evidence* (ed. Delson, E.) 14 (Alan R. Liss, New York, 1985).
18. Raza, S.M. et al. *Nature* 306, 52 (1983).
19. Sankhyan, A.R. *J. hum. Evol.* 14, 573 (1985).
20. Wu Rukang & Oxnard, C.E. *Nature* 306, 258 (1983).
21. Wu Rukang, Lu Qiuwu & Xu Qinghua. *Acta Anthropol. Sin.* 3, 1 (1984).
22. Martin, L. *Nature* 314, 260 (1985).
23. DeBonis, L. & Melentis, J. C. r. *Hebd. Séanc. Acad. Sci., Paris. Ser. II* 300, 429 (1985).
24. Martin, L. & Andrews, P. *Cour. Forschungsinst. Senckenberg* 69, 25 (1984).
25. Kelley, J. & Pilbeam, D.R. (abstract) *Am. J. phys. Anthropol.* 66, 188 (1985).
26. Pilbeam, D.R., Rose, M.D., Badgley, C. & Lipschutz, B. *Postilla* 181, 1 (1980).
27. Pilbeam, D.R. *Bull. Yale Peabody Mus. Nat. Hist.* 31, 1 (1969).

Eric Delson is at the American Museum of Natural History and the Department of Anthropology, Lehman College, City University of New York, Bronx, New York 10468, USA.

Keeping science at the core

Harvey Brooks

The Education of a College President: A Memoir. By James R. Killian, Jr.
MIT Press: 1985. Pp.478. \$19.95, £19.95.

DURING the middle years of this century, James Killian presided over the transformation of the Massachusetts Institute of Technology from a highly respected engineering school to the world's leading technological university, a "university polarized around science". This last phrase is used by the author to describe an institution embracing virtually all fields of knowledge, but with each discipline in orbit, as it were, around a central nucleus of science and technology. In the same career, Killian intermingled university leadership with high-level public service, notably as science adviser to President Eisenhower and as a key figure in many seminal government advisory groups during a formative period for both science policy and defence policy. What were the experiences and personal qualities that enabled this individual, not himself either a scientist or an intellectual with academic pretensions, to become accepted as a colleague and leader by scientists and scholars, and to become an educational and scientific manager of the first rank? This highly personal, rambling and optimistic memoir does provide some insight into the answers, though it is as interesting for what it leaves out as for what it includes.

A man of broad general culture, boundless intellectual curiosity and catholic interests, very "people oriented" though not remarkably gregarious, Killian became dedicated through an unusually varied apprenticeship to "developing creative integration and interdisciplinary congeniality among a variety of research fields", and he applied an extraordinary talent for consensus-building among diverse and strong-minded individuals in both academic and political settings. He was obviously a good listener with a capacity for extracting a practical synthesis out of the ideas of creative people. Not an obvious originator himself, and lacking a strong pride in authorship, he nonetheless had a talent for recognizing people with original ideas and enabling them to realize their visions.

His success in these endeavours was reinforced by what he describes as his "clinging to a melioristic view that our future is best served by an acceptance of what Sir

Peter Medawar has aptly termed 'the hope of progress' ". This attitude, he felt, came from his "having lived so long in the regenerative sanctuary of a research university" with a "singular fellowship of young men and women working in a contagious atmosphere of excellence, discovery, and high spirits" (p. 415). He had a knack for



James Killian on the cover of *Newsweek*, November 18 1957.

helping brilliant — often cranky and egotistical — people to sell themselves and their ideas, and he believed this could best be accomplished through conscientious and patient consultation.

Insofar as MIT was concerned, Killian rejects Clark Kerr's pessimistic view of the modern "multiversity" impervious to structural change, the advocate of change for every part of society except itself. The MIT he knew originated and assimilated many structural improvements, including new interdisciplinary laboratories, curriculum reform, undergraduate research, a new management school, new departments in the humanities and social sciences, a joint school of health sciences and technology with Harvard, and many other institutional innovations. MIT, Killian feels, has avoided the usual "hang-ups" of the multiversity by combining a single unified faculty with a strong, but accessible, style of presidential leadership, able to

persuade the disciplines to subordinate their parochial interests and claims to academic "turf". Thus he sees MIT as a federal rather than as a feudal system. I suspect one could find a good many people at MIT who would not fully share this view, but I think it is, on average, correct.

Killian cites two revealing anecdotes to illustrate the source of his view of the key importance of consensus-building. Each involved the eventual failure or compromise of a good idea because of a lack of consultation with powerful interests in advance of a public commitment. The first example is an academic one. In 1956, at a public dinner in New York, he announced MIT's intention to establish within its structure a new School of Advanced Study before he had spoken to the appropriate faculty committees or the faculty as a whole. From the incident he drew a lasting lesson that "no new programs or buildings large in costs should be undertaken, however innovative, without the knowledge and advice of appropriate faculty groups" (p.138). In this instance, what he still obviously feels was an excellent idea had to be abandoned.

A parallel example on the national stage occurred soon after Killian had become science adviser to Eisenhower. In 1958, on the basis of the technical findings of the President's Science Advisory Committee, communicated to him by Killian, Eisenhower wrote to Khrushchev suggesting that a comprehensive nuclear test ban agreement was negotiable, and proposing the convening of a panel of technical experts to agree upon the design of an international detection system capable of ensuring against clandestine testing — especially underground testing — by any nation. The President's initiative, however, was taken without internal consultation with the Department of Defense, The Atomic Energy Commission or the Central Intelligence Agency.

Of the technical discussions, led on the American side by Dr James B. Fisk, President of Bell Laboratories, Killian says: "Dr. Fisk and I agreed that at no subsequent negotiation with the Soviets on arms limitation had the Soviets expressed such willingness to accept on-site inspection" (p.281). He implies that the lack of internal consultation in the US government before the Presidential initiative "was to aggravate differences in our own government" which became a major factor in America's inability to negotiate a comprehensive test ban rather than the much more limited atmospheric test ban finally achieved under the Kennedy administration in 1963. As we now know, the

rapid evolution of a sophisticated technology for underground tests largely offset the constraints on weapons development posed by the ban on atmospheric testing, and really only served to drive nuclear weapons development underground — literally and figuratively — and out of the public eye. Whether prior consultation would have merely tied the hands of the technical negotiators rather than leading to a stronger diplomatic agreement is a matter of opinion. But again, Killian's account of this episode illustrates the high importance which he attached to the building of consensus among divergent interests as a prerequisite for action (p. 279).

Although Killian's assessment of the achievements of the modern research university is highly positive, and in many

the quality of engineering and is this in turn a consequence of deterioration in the standards of engineering education? Were the reforms of the engineering curriculum to make it more "scientific" and general, which Killian obviously endorses, implicated in the problem, as some people are now asserting? Did the American preoccupation with originality, creativity and individual competitive achievement somehow result in neglect of meticulous workmanship and attention to the more routine and mundane aspects of engineering, which are so crucial to the quality and reliability of technological products? Did the great research universities, in their pursuit of a particular cultural concept of "excellence", provide an inappropriate model for the mass of the engineering and technical professions?



James Killian (second from right) with Philip Morse, Vannevar Bush and Thomas J. Watson dedicate MIT's first major computational centre in 1957.

ways forms the *leitmotif* for his whole memoir, he does concede some shortfalls and failures; however, he makes little effort to analyse the possible reasons for them. For example, while he is proud of MIT's success in creating interdisciplinary laboratories and in fostering other kinds of creative work across conventional intellectual boundaries, he admits that MIT's efforts "have not succeeded in demonstrating how these fields necessarily interact with each other and how this diverse learning is comprehended in indivisibility" (p. 402). Somehow his "enlarged vision of the place" has not quite been achieved. Yet he never even speculates as to the reasons why this might have been so.

More seriously, he suggests that the universities cannot wholly disclaim responsibility for an unmistakable decline in the quality of everyday life as so frequently reflected in American craftsmanship, in the unreliability of too much of our technology, and in the low quality of many consumer products and their service [p. 410].

This is an intriguing statement, which is simply left hanging. In what way or in what sense are the universities responsible? Does the situation reflect a decline in

Another intriguing area of ambivalence in Killian's account of the era he describes has to do with the role of the "military-industrial complex". On the one hand, the whole memoir is full of pride in the services of MIT and its staff (including Killian himself) to the nation in matters of national security — the series of interdisciplinary studies organized by MIT in the 1950s, the creation of the Lincoln Laboratory, the accomplishments of the Instrumentation Laboratory (now divested from MIT as the Draper Laboratory), Project Whirlwind, the Servomechanisms Laboratory and so on. On the other hand, he makes frequent references to the dangers of militarization, stating in one place that "the military-industrial complex . . . is still having a profoundly dangerous impact on our government" (p. 419). Even his "hope of progress" is "tempered . . . by fear of nuclear war and of an intemperate militaristic spirit eroding the most precious values of our society" (p. 415).

These are strong words, but just who is the military-industrial complex? Many of the basic technological advances that have helped fuel the arms race have emerged from the large government-financed laboratories managed and operated by some

of our great research universities — for example, the concept of multiple independently targeted re-entry vehicles and dramatic advances in missile guidance accuracy that emerged from the MIT Instrumentation Laboratory, and the high yield-to-weight thermonuclear weapons which were developed in the two weapons laboratories of the University of California. While praising the scholars from research universities who "have been speaking with knowledge, integrity, and responsibility about nuclear issues" (p. 419), Killian never quite comes to grips with what the proper role of the research universities should be in relation to the creation of military technology or even the science underlying it. Should research universities, for example, pick and choose which governmental demands they respond to according to some collective view of their potential effect upon the arms race? What would be the implications for academic freedom of the enforcement of such a collective sense of responsibility on their individual members? To what extent would such institutional self-denial decrease the capacity of academics to speak out with real knowledge and responsible criticism in the national debates over defence policy, and leave the field to those with greater vested interests than the universities? These issues have never been posed to the universities in starker form than by the current Strategic Defence Initiative; it would have been interesting to have had Killian's views as to how his favourite institutions should come to grips with this new challenge to their role.

The sanguine tone of this memoir sometimes borders on the polyanna-ish. It is striking that, although there are interesting personal sketches of many of the principal actors in the development of America's post-war science and defence policies, villains are notable by their absence. Admiral Lewis Strauss of the Atomic Energy Commission is virtually the only important personality even mentioned in unfavourable terms, and even here the strictures are remarkably restrained. The vehement denunciations of the military-industrial complex are notable for their lack of specificity, and the reader is left in the dark as to which institutions and groups qualify for this pejorative label, and what sorts of behaviour take them over the line from patriotism to villainy. Lobbying of the political process on behalf of funding for particular weapons systems seems to be the only identified behaviour which is beyond the pale, yet one must also ask how these activities are to be distinguished from the various "summer studies" and high-level commissions (such as the Technological Capabilities Panel, which Killian himself chaired) which were the genesis of many new weapons systems.

Harvey Brooks is Benjamin Peirce Professor of Technology and Public Policy, John F. Kennedy School of Government, Harvard University, Cambridge, Massachusetts 02138, USA.

Chatting fractals and New Math

Thomas Banchoff

Mathematical People: Profiles and Interviews. Edited by Donald J. Albers and G.L. Alexanderson. *Birkhäuser*: 1985. Pp.372. \$24.95.

GEORGE Pólya died recently at the age of 97. Here was a legendary teacher whose book *How to Solve It* and famous film *Let Us Teach Guessing* inspired generations of students and teachers of mathematics. Would it not have been instructive to have had the opportunity to ask him questions about his career? In recent years, Benoit Mandelbrot's filigreed fractal patterns have burst out of the pages of journal after journal, intriguing scientist and designer alike. How would it be to interview him and get some insight into the process of creating a new kind of mathematics?

In *Mathematical People* we can enjoy conversations with these figures and two dozen more. This collection of interviews and profiles comes primarily from a very successful series of articles over the past six years in *The College Mathematics Journal* (formerly *The Two-Year College Mathematics Journal*). The project began auspiciously with an interview with Pólya on the occasion of his ninetieth birthday celebration at Stanford and continued to include some of the most interesting mathematical people of our day.

Here are stories about Persi Diaconis, who left school at 14 to travel with a professional magician, Martin Gardner who explains the origin of his hugely popular *Scientific American* column, and the immensely prolific, champion collaborator Paul Erdős. We can follow the controversial views of Morris Kline about the New Math, and compare them with the counter-suggestions of Peter Hilton and Paul Halmos. We can contrast the impressions of Richard Courant given by his co-author Herbert Robbins and by the biographer Constance Reid. And we can keep track of dozens of comments about the sometimes opposing relationships between pure and applied mathematics.

What we do not find are answers to the question: "What good is mathematics anyway?" The interviewers are for the most part mathematicians themselves and everyone takes for granted the fact that the creation and communication of mathematics is a worthwhile occupation. Moreover the primary readership for the original articles was teachers of mathematics, so questions tend to centre on different aspects of teaching as opposed to research.

There is wide variation in the depth and the nature of the interviews and profiles, and occasionally a general reader will be left behind when the discussion takes a

technical turn. There are also a few other places where one looks in vain for a follow-up to a remark in an interview. For example, what was the subject of Donald Knuth's first published article, in *MAD Magazine*? Could Albert Tucker please explain why his mentor Solomon Lefschetz had artificial hands? And if David Blackman was not conscious of discrimination all through his schooling, why was it that he only applied to black institutions when he sought his first job?

By and large, however, the interview format is quite successful. By contrast the set pieces taken from popular scientific journals, on Shing-Shen Chern and Ron

Graham for example, may be better organized and generally more informative but they lack the spontaneity of the transcribed interviews. The detailed and technical autobiographical reminiscence of Olga Taussky-Todd, originally written as part of a Caltech oral history project, seems out of place in this volume.

Still, one must expect such vagaries in a collection such as this. Many people will find illumination and entertainment in the book, and will be happy to learn that a second volume is in preparation. □

Thomas Banchoff is Professor of Mathematics at Brown University, Providence, Rhode Island 02912, USA.

In physical context

Ivor Grattan-Guinness

Mathematics and the Search for Knowledge. By Morris Kline. *Oxford University Press*: 1985. Pp.257. \$19.95, £21.

OVER several decades, Morris Kline has built up an enviable reputation as the author of a series of popular books on mathematics which include a substantial element of historical material. The first book, *Mathematics in Western Culture* (Oxford University Press, 1953), remains perhaps the best: the last one, *Mathematics: The Loss of Certainty* (Oxford University Press, 1980), was probably the worst, with numerous infelicities and errors over the foundations of mathematics and its history. This latest volume lies somewhere between the two.

Here, Kline has returned to applied mathematics, in order to explain how mathematics helps us to understand the physical world. After a "Historical Overview" dealing with various philosophers' positions on the possible existence of an external world, and a discussion of the fallibility of sense perception, he proceeds to a chronological account of episodes in the history of astronomy and mechanics: the Greeks, Copernicus and Kepler, Descartes and Newton. Then electromagnetism is introduced, chiefly as background to relativity and quantum mechanics. The final trio of chapters contains discussions of aspects of the philosophical questions which underpin the book as a whole.

The level of success achieved by Kline in answering these questions may be appraised at first by looking at his use of history. The treatment is more detailed than in his *Mathematics and the Physical World* (Crowell, 1959), but not much; in particular, there is not actually a great deal of mathematics in the book, the chapters on general relativity and quantum mechanics including hardly any at all. Further, the range of applications is curiously incomplete: there is little or nothing on optics, heat diffusion (the paragraph on p. 202 about Fourier mis-

represents his philosophy, overrates his mathematics and underrates his physics), hydrodynamics, elasticity theory or engineering mathematics (friction studies, machines of all kinds, cartography and topography, and so on). Probability and statistics earn the falsehood that "the theory of probability . . . entered into mathematics quite by chance in connection with games of chance" (p. 189): the virtual absence of these two topics is an especial pity, since work on them greatly extended the range of application of mathematics and raised important questions about knowledge and ignorance, and the epistemological status of approximate theories.

The result of these omissions is not only



Vaulting Ambition

Sociobiology and the Quest for Human Nature

PHILIP KITCHER

'The best dissection ever published of the logic and illogic (mostly the latter) of sociobiology.'

Stephen Jay Gould

'Vaulting Ambition is indeed the last word on the subject of sociobiology. It is meticulous in its argument and total in its scope. It deals a carefully reasoned blow to the pretensions of sociobiologists.'

Richard Lewontin
December £24.95 Hardback
480pp 0-262-11109-8

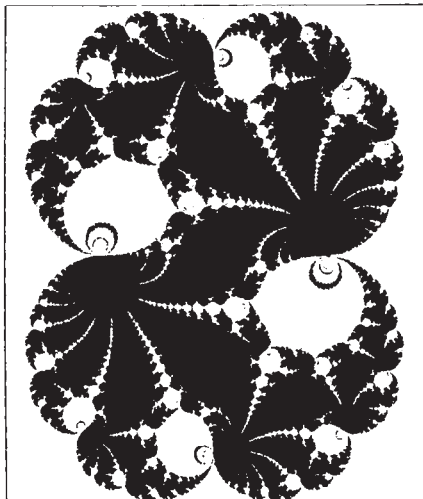
THE MIT PRESS

126 Buckingham Palace Road, London SW1W9SD



Reader Service No.21

that many essential historical episodes have been passed over, but also that certain crucial aspects of the ways in which mathematics aids our search for knowledge are absent. The terms within which the subject has been tackled are too modest, a failing which also applies to the weight of supporting mathematical apparatus which is brought to bear. There is, for example, no discussion of the calculus, so that the "mathematicization" of continuous media and phenomena cannot be described properly. The clash between different mathematical styles is not



A Julia set and its interior for the map $z \rightarrow z + iq$. Taken from *Mathematical People* (see previous review).

brought out, especially between the geometrical (where the properties of the given configuration play a role) and the algebraic (where formal manipulations of general basic formulae are executed, with deliberate avoidance of geometrical processes); analytical and topological styles bring further differences in their train. Finally, the extent to which a mathematical expression (in any style) reflects the structure of the physical context in which it is used is a matter which should have loomed large, but it does not.

All of these questions involve issues about mathematical proofs, which receive little attention here. But proofs lie at the heart of much of the development of mathematics and thus of its role in the "search for knowledge". The posing of such questions would have served Kline's cause better than the pastiche accounts of philosophers' views on the external world which he does provide (but which have no special bearing on the use of mathematics). This book is best taken as another example of his ability to write breezy and readable surveys of the history of mathematics, and to tackle the philosophical background on the wing, as a contribution to the problem raised, however. *Mathematics and the Search for Knowledge* is merely a propaedeutic. □

Ivor Grattan-Guinness is Reader in Mathematics, Middlesex Polytechnic, Enfield, Middlesex EN3 4SF, UK.

Where lies the science?

Anthony W. Clare

The Psychoanalytic Movement. By Ernest Gellner. *Paladin: 1985. Pp.241. Pbk £3.50.*

Psychoanalysis and Beyond. By Charles Rycroft. Edited by Peter Fuller. *Chatto & Windus: 1985. Pp.310. Hbk £10.95; pbk £4.95. To be published early next year in the United States by Chicago University Press.*

Psychoanalysis: Freud's Cognitive Psychology. By Matthew Hugh Erdelyi. *W.H. Freeman: 1985. Pp.303. Hbk \$24.95, £25; pbk \$14.95, £14.95.*

Decline and Fall of the Freudian Empire. By Hans Eysenck. *Viking: 1985. Pp.224. £12.95.*

THE way in which the psychoanalytic school was formed raises profound questions concerning its status as a science and suggests that it might, with benefit, be viewed as a secular religion. Certainly, one will look in vain elsewhere in science for another example of a body of theoretical knowledge linked so intrinsically, so inextricably, with the individual who first propounded it. In no other branch can one find an absence of an external arbitrator, that is to say someone outside the control of the system itself and able to pass independent comment on its claims, achievements and shortcomings. In the case of psychoanalysis, the location of the criterion of its truth is, in the words of Ernest Gellner, "determined by the ideas of the system itself; the well of truth is within the ramparts and not outside". The possession of this truth is not handed down in seminars and courses of instruction, though these play a small part, but by way of participation in the training analysis. As the apostolic succession from St Peter is handed down from the bishops to newly ordained priests, so the authentic link with Freud is maintained and continued by means of the transfer of authority from training analyst to analysand.

The religious analogy does not end there. As intellectual and rational criticism of revealed religious dogma is interpreted as a closing of the mind and heart to truth, and in no way a valid comment on the basic validity of the truth itself, so criticism of psychoanalysis is rejected on the grounds that either the critic has never been analysed, in which case he is resisting the need to be analysed (which itself is worthy of interpretation), or he has been analysed and his criticisms reflect an incomplete or prematurely terminated process, the result no doubt of fear over what might be revealed. And the religious analogy becomes uncomfortable when one examines the schisms and factions within the domain of psychoanalysis, with each school claiming to have adhered the

closest to the spirit as well as the law laid down by the supreme founder.

The question which Gibbon asked of Christianity is equally applicable to psychoanalysis: by what means did the new vision obtain so remarkable a victory? In a stylish, witty and deceptively readable book, Gellner exposes the secular religious nature of the psychoanalytic enterprise. He admits that a compelling, charismatic belief must possess more than merely the promise of succour in a plague and links with the background convictions of the age. It must engender a tension in the neophyte or potential convert. It must tease and worry him with its promise and its threat. The doctrine to be embraced must contain the promise of a salvation that is earnestly desired. There must be plausible reasons for believing its claims and also good reasons for doubting it or fearing its truth. The middle ground between acceptance and rejection must be denied and the issue of whether it is or is not true must be trans-empirical and untestable. The fact that confidence in psychoanalysis is devoid of any visible means of support hardly seems to diminish its authority: indeed, argues Gellner,

were its means of support visible, they would be subject to evaluation; and the assertion would have given hostages to fortune, and also conceded its own ordinary, human status — making its own contentions arguable, negotiable, imitable, publicly available rather than tied to a single source.

One psychoanalyst who has been particularly conscious of the contentiousness of the psychoanalytical factions is Charles Rycroft who, in so far as he identifies with any particular school, tends to align himself with the English-speaking school uniting such disparate practitioners as Winnicott, Fairbairn, Balint and Bowlby. This group, as Rycroft pointed out in an essay entitled "Psychoanalysis and Beyond", contained in the collection of his writings under the same title, repudiated any royal road to becoming an analyst even to the extent of encouraging trainees to have part of their teaching under the supervision of a member of another faction. Rycroft underwent, as far as one can tell, an orthodox Freudian analysis and in 1956 presented a paper to the British Psycho-Analytical Society which challenged the prevailing theory of symbolism by suggesting that it is not a regressive or defensive phenomenon but rather "a general capacity of the mind" which could be deployed in manifold different ways. According to Peter Fuller, the editor of the book, the Society received this modification, the first serious one advanced to a theory of some 40 years standing, with reserve while it struggled to protect the orthodox position. It was the first of a number of theoretically awkward questions that Rycroft was to raise, on such issues as dreams, fantasies and the nature of psychoanalysis itself.

In 1968, he wrote *A Critical Dictionary*

of *Psychoanalysis*, a task which enabled him to take all the current psychoanalytical theories and systems to pieces and compelled him to construct some kind of overview embracing the ideas of Freud and the British analysts — such as Anna Freud, Melanie Klein, Michael Balint and Donald Winnicott — who had been active when he himself had been a student. His growing conviction that psychoanalysis is not so much concerned with the causes of human behaviour as with its meaning, and is better seen as one of the humanities than a science, does enable him to side-step many of the objections to psychoanalysis made by those such as Gellner and, as we shall see, Eysenck. However, nowhere in the 25 essays contained in this collection does he apply his mind to the issue which clearly troubles critics, namely the extent to which psychoanalysis is a self-fulfilling explanatory theory, whether it be of meaning or causes.

Meanwhile, there are others such as Professor Erdelyi, of the Department of Psychology at Brooklyn College, who struggle to salvage as much of Freud's expansive theory as modern insights in cognitive, physiological and neurological psychology will permit. Erdelyi's book uses case histories, laboratory studies, newspaper cuttings, visual art and poetry in an imaginative if occasionally wayward examination of Freudian theory. What emerges, however, suggests that what was lasting in Freud's conceptions is their general preoccupations. The specific elements of that gigantic canvas — the separation of mental functioning into id, ego, superego, theories of dream function and the nature of repression — just do not stand up to the relentless scrutiny of the laboratory scientists or the critical clinician. A particularly good example concerns Freud's pleasure principle. Freud assumed that pleasure/positive reinforcement resulted from tension-drive reduction; that, in short, pleasure was the cessation of pain — the so-called "Nirvana principle". The physiological facts suggest

otherwise. Pleasure and unpleasure appear to be not opposite sides of the same coin but two different coins altogether. More uncomfortable for psychoanalysts is the fact that the "pleasure" arising from stimulation of the pleasure centre in the brain is general in nature and *not* sexual as a great deal of psychoanalytical writing and excavation argues.

The most fundamental criticism of psychoanalysis, however, remains the same, 30 years on from when Hans Eysenck first mounted it. It is the fact that there is no convincing evidence that patients do better under analyses than they do under some other form of psychotherapy or psychiatric treatment. Such criticism does not come only from without. Anthony Storr, himself a Jungian analyst, has concluded that "the evidence that psychoanalysis cures anyone of anything is so shaky as to be practically non-existent". And Brian Farrell, a philosopher by no means hostile to analysis, has admitted that the impact of psychoanalysis

cannot be justified on the grounds that it contains a body of reasonably secure or established knowledge . . . it [is] only too painfully evident that analysis does not contain any such body of knowledge.

Characteristically, Eysenck puts the case with more exuberance and venom. At best, he concludes his devastating polemical assault, psychoanalysis is "a premature crystallization of spurious orthodoxies; at worst a pseudo-scientific doctrine that has done untold harm to psychology and psychiatry alike". The final verdict, however, is not in and it is doubtful that it will be within the lifetimes of the authors of these books. What does seem unarguable, however, is that in the one hundred years that have passed since Freud began to build his theoretical framework the theory still lacks any substantial validity and his school remains beleaguered. □

Anthony W. Clare is Professor in the Department of Psychological Medicine, St Bartholomew's Hospital Medical College, West Smithfield, London EC1A 7BE, UK.

Scent of semi-science

Steve Blinkhorn

The Intelligence Men: Makers of the IQ Controversy. By Richard E. Fancher. W. W. Norton: 1985. Pp. 269. \$17.95, £14.95.

AT LAST, it seems, the soot and whitewash school of controversy has had its day. Here is a book about the IQ debate which is at once absorbing, informative, deftly written and, if the expression can be pardoned, a thundering good read.

The Intelligence Men is that rare inversion of the natural order of things: a first-rate book based on a second-rate notion. The notion, that one may discover why particular figures in the history of a debate took up their respective positions by examining their personal histories, is only as good as the sample of biographical material available and the accuracy with which it is interpreted; in any case, the standards of proof required of such a notion are far inferior to those proper to the debate itself. Fortunately analysis of this sort intrudes very little into Fancher's account of the lives and work of a dozen or so major contributors to the long-running dispute over IQ.

The delicious consequence of this plan has been a book reeking of the dubious scent of semi-science in the making. Where often a scientist is known to a wider public only as an author of published work, here that work is set in a personal context which illuminates and informs even when the technical level of the narrative is less than satisfactory. Pasteboard textbook characters have pasteboard theories to peddle. But the facts that Spearman registered for a PhD with Wundt (and took seven years to complete it), that Terman's introduction to psychology was a book on phrenology and that amongst Yerkes's motivations for seeking a career as a physician was the fact that it paid better than being a farmer, bring an unusual sparkle to academic history.

Kamin's brush with McCarthy's subcommittee on unAmerican activities; Jensen's origins as the son of a lumber and building-supplies dealer, and habit of conducting recorded symphonies at home with a chopstick for a baton; Binet as a child being forced by his father to touch a cadaver; Burt labouring to devise raw data to fit published statistics — none of this is relevant to the logic of scientific discovery. But, as Fancher suggests, the psychology of the scientist is at least as potent a force in determining the direction of theory. This, of course, is a view that can be taken to extremes. There was once a claim that Spearman, being British and therefore a monarchist, was driven to expect a single factor of intelligence, whereas Thurstone, being American and more democratically inclined, preferred a federalist view with several distinct abili-

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Sigmund Freud — still dominant in the field of psychoanalysis.



On the track of Ice Age Mammals

A.J. Sutcliffe

This fascinating and extensive survey of the Ice Age, taken from a multi-disciplinary standpoint, will become a mine of information both for the research scientist and the layman. Fully illustrated throughout, featuring 5 double-page colour spreads commissioned from artist Peter Snowball.

224pp, 150 b/w illustrations + 5

double pages colour.

0 565 00869 2. Hbk. £12.95

Crows of the World, 2nd ed.

Derek Goodwin

An up-dated edition of this standard reference work on the familiar yet diverse crow family. Includes 3 new colour plates commissioned from Robert Gillmor.

300pp, approx. 6 colour plates. 1986.

0 565 00979 6. Hbk, approx. £30.00

The Littoraria species of Indo-Pacific mangrove forests.

D.G. Reid

A detailed study of 20 species of *Littorina scabra*, many previously unrecognised, with particular attention to the reproductive system by which the species can be reliably distinguished. A key to shells is also provided.

240pp, 99 illustrations, 1 colour plate.

0 565 00978 8. £35.00

A world list of mammalian species, 2nd ed.

G.B. Corbet & J.E. Hill

New edition of this easy-to-use reference book designed primarily for the non-specialist. Incorporates many additions and revisions since the first edition of 1980.

240pp, approx. 1986.

0 565 00988 5 £25.00

Publications Sales, British Museum, (Natural History), Cromwell Road, London SW7 5BD

ties ranking equally. Nothing so crude enters Fancher's book; rather, there is an elusive subtlety to the drift of the narrative that no review can hope to capture.

It is all the more unfortunate, then, that the last chapter reads like a hurried afterthought. In attempting to come to a balanced view of the current state of the debate, the author steps outside his role of historian and immediately becomes unconvincing. This is a pity, but a small blemish. As the origins of intelligence testing recede beyond living memory, this book does a real service by reminding us that present academic fisticuffs have their origins in an honest but often naive and blinkered struggle to make sense of human variability. No doubt future historians will judge present efforts to have been equally blinkered: but in what proportions are we now naive and honest?

Steve Blinkhorn is Director of the Psychometric Research Unit and Principal Lecturer in the Department of Psychology, Hatfield Polytechnic, PO Box 109, College Lane, Hatfield, Hertfordshire AL10 9AB, UK.

Go for Valhalla

Michael Spencer

The Joy of Science: Excellence and its Rewards. By Carl Sindermann. Plenum: 1985. Pp.259. \$16.95, £16.15.

SAMUEL Smiles, the Victorian author of such uplifting texts as *Self-Help* and *Men of Invention and Industry*, has been enjoying a revival of esteem on the New Right. Britain's Prime Minister has told unemployed teenagers to go out and start their own businesses, while one of her ministers has coined the immortal phrase "Get on yer Bike". We are urged to model ourselves on those who pursue success with relentless enterprise and determination.

Carl J. Sindermann is the Samuel Smiles of science, and has followed his earlier book *Winning the Games Scientists Play* with a further guide on How to Get On. His technique is modelled on that of Smiles: to examine in detail the upward climb of those who have achieved success. By "success" he means not the winning of Nobel prizes (a chimera that motivates more scientists than would ever admit it), but a job-satisfaction extending to administration and the manipulation of grant-giving bodies.

He cheerfully admits in his preface that the prospect of such a book did not endear him to his colleagues, who may have suspected that their own careers (thinly disguised under fictitious names) would be dissected for all to see. Undaunted, at late evening cocktail parties and professional meetings he pursued his chosen paragons long after their normal bedtimes, so as to analyse what made them tick.

The chapter on "The Ascendant Female Scientist" was, says the author, written against the strongly stated advice of all his female acquaintances, who were still smarting from a section entitled "Sex in the Laboratory" in his previous book. None of them, he says, would admit to the classification of "friend" after it was published.

Why, then he do it? The answer appears to be that Dr Sindermann is an old-fashioned scientific optimist who believes that all science is exciting and wonderful, and that anyone who succeeds in it is to be applauded and emulated. (In this he may have something in common with the august figure who once opined in print that it would be a pity for the world to blow itself up with nuclear weapons because that would halt the onward march of science.) Hardly anywhere in the book — whether discussing choice of research topic, career transitions or interactions with the outside world — does he discuss the problem of deciding whether a project is socially desirable. The word "ethics" does not appear in the index.

Dr Sindermann has little time for the also-rans, except as examples of where one can go wrong. He would certainly not agree with a friend of mine, who emerged limply from a two-hour session with a visiting celebrity to remark that the failures were much more interesting to talk to. He has harsh words to say about the Burnout, the Fade-out and the Guaranteed Loser. The latter category includes the Unwary Activist, who may (he says) damage his or her scientific credibility by espousing causes for social reform. The author classes such misguided people along with the Dilettante, the Hobbyist and the Chronic Underachiever.

Since most of us (by definition of that uncompromising word "excellence") are not going to reach Dr Sindermann's Valhalla of success, why should anyone read his book? The upwardly mobile graduate student with a taste for ruthless self-advancement will have his own ideas on how to make it. For those who have reached what the author calls the "midlife crisis" it is probably too late. The ageing scientist (to whom a chapter is devoted) is not going to believe a word of it.

Perhaps the most interesting thing about the book is what is missing from it. There is the same moral vacuum in the Royal Society's message to scientists, that they should learn how to manage the media in support of their chosen way of life. It is possible — just possible — that the current public mood of suspicion and alarm about the outcome of some scientific research is due not to inadequate public relations, but to a failure among scientists to think about the repercussions on ordinary people. Politicians have made the same mistake. □

Michael Spencer is a voluntary exile from science based at 34 Bayham Road, Sevenoaks, Kent TN13 3XE, UK.

Modularity of brain and mind

P.N. Johnson-Laird

The Social Brain: Discovering the Networks of the Mind. By Michael S. Gazzaniga. Basic Books:1985. Pp.240. \$17.95.

Neuronal Man: The Biology of Mind. By Jean-Pierre Changeux. Translated by Laurence Garey. Pantheon:1985. Pp.348. \$19.95.

THE brain is the organ of the mind. If you want to understand the mind, then you had better understand the brain. These two books, not perhaps as different as chalk and cheese, but certainly as different as brie and parmesan, share this philosophy. They both aim to elucidate the mind by establishing how the brain works. Michael Gazzaniga wants to demonstrate a link between the global organization of your brain and the way in which you come to believe in certain propositions — from prosaic assumptions about yourself to your propensity to hold religious beliefs. Jean-Pierre Changeux is a Gallic ghostbuster, who aims to rid you of your faith in the "ghost in the machine". He will explain your mind away: it is nothing but neurones, synapses, and electrical and chemical signals. His motto is La Mettrie's: "The soul is merely a vain term of which we have no idea. Let us conclude boldly that man is a machine".

Gazzaniga is a distinguished neuropsychologist at the Cornell University Medical Center in New York, and is best known for his studies of the psychological consequences of the "split-brain" operation. His book is an intellectual autobiography that recounts the development of these studies, considers their implications and speculates about such larger issues as society, prehistory and religion. He writes in a loose, unbuttoned style, with numerous asides and anecdotes, and is at his strongest in describing his beautiful series of experiments on the effects of the split-brain operation.

If you fixate some point in the scene in front of you, then everything to the left of your fixation point is projected to the right half of your brain, and everything to the right of your fixation point is projected to the left half. Information is rapidly transferred from one half to the other by a massive bundle of nerve fibres, the corpus callosum. The split-brain operation is carried out to control extreme forms of epilepsy. The surgeon cuts the corpus callosum and its associated structures so as to isolate the left cerebral cortex from the right. The operation is highly beneficial in otherwise intractable cases, and, according to early results, had no effects on personality, mood or behaviour. However the absence of behavioural consequences did not square with some earlier studies on

animals carried out by Roger Sperry, and so Gazzaniga decided to check again. Under Sperry's supervision, he carried out the pioneering series of studies that revealed the operation's consequences.

The first effect he observed was the patients' inability to name anything that lies in the left visual field: such a stimulus projects to the right hemisphere, but the main language centres are located in the left hemisphere. Information can no longer be transmitted from one hemisphere to the other, and so the patients deny seeing the stimulus. However, when they are asked to make guesses about it, their emotional evaluations are almost identical to those they make when they can see it. They become frightened by terrifying pictures, though they still deny seeing them, and may attribute the cause

such distractions and thus scores higher on perceptual tasks.

When split-brain patients are given two matching tasks, one to solve in the left brain and the other to solve in the right brain, they readily perform them, but the verbal justifications they give for the responses made by the right brain are usually false even though the responses themselves are correct. Thus, if the command "walk" is flashed to the right brain, then the patient will indeed get up and start to walk. If asked why, the left brain, which is unaware of the command, readily concocts a confabulatory explanation. Here is the seed for Gazzaniga's principal hypothesis.

The brain is organized, he claims, into separate modules that can carry out various actions, maintain moods and per-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Wellcome Institute

An early look at the raw material — structure of part of the brain from Julius Casserius *Tabulae anatomicae*, Frankfurt 1632. Book x, pl. iii.

of their fear to the demeanour of the experimenter.

If I remove the cooling fan from my micro-computer, the machine is soon incapable of carrying out any computations. Someone who knows little about computers might therefore infer that the fan plays a central role in computation. Analogous errors are an ever-present danger in the study of brain damage. Gazzaniga wisely resists them and repudiates the one-time vogue for alleged dichotomies between left-brain thinking (analytical and verbal) and right-brain thinking (intuitive and visual). His further studies have revealed some subtle phenomena. The right hemisphere typically can create better drawings than the left. Yet the right is not necessarily a superior visualizer: it does no better than the left in matching one readily nameable picture to another. Gazzaniga suggests that perhaps the right brain does not possess a special ability at perception. Instead, if the left brain has no immediate opportunity to exercise its linguistic ability, it may concentrate still more on the verbal aspects of the task — to the detriment of its perceptual performance — whereas the right hemisphere has no

ceive the world. They are independent units that work in parallel; and they contribute to what we are conscious of, but we are not aware of them. They can control behaviour, and so we sometimes do things for no reason that we are aware of — we act capriciously. The module that mediates language and consciousness, however, interprets our behaviour, and instantly contrives theories to explain it. That is why we think we have free will, and why we are likely to hold beliefs for which we have no objective evidence.

Another inherent danger in interpreting brain damage is to overlook the possibility that its effects are a result of compensatory changes. A sceptic might accordingly argue that Gazzaniga is right: one half of a split brain does contrive explanations for how the other half lives, but only because it has been forced to do so as a result of the operation. Normal people behave quite differently and are conscious of the roots of their behaviour. Of course, there is a long tradition in psychology — and considerable evidence — to the contrary. Likewise, the modular hypothesis has been independently postulated by a number of cognitive scientists in order to

make sense of the mechanisms of vision and language, the distributions of scores on different sorts of mental test, and other neurological disturbances. What is unique in Gazzaniga's account, however, is his emphasis on the confabulatory capacity of the interpretative module.

Knowledge of the brain has changed strikingly in the past 25 years. It used to be assumed that sensory information was conducted to the cortex, where there was an interchange with other information in the so-called "association" areas, and finally neuronal signals were transmitted to the motor areas responsible for the control of actions. This picture is now known to be a vast oversimplification. Jean-Pierre Changeux, who is a molecular neurobiologist at the Collège de France, has made important contributions to these developments in his work on neurotransmitters and the growth of neural networks. His prize-winning book was evidently a best-seller in France, and it is an outstanding attempt to convey to the general public an interdisciplinary understanding of the human nervous system.

Changeux begins with the tortuous history of concepts of the brain. He records the gradual recovery from Aristotle's original blunder in assuming that it functioned as a cooling device — a case of confusing the central processor with the fan — and describes the gradual ousting of the idea of "vital forces" by the modern electrochemical theory of the propagation of nerve impulses. The brain contains a dense interlacing of the dendritic webs of thousands of millions of neurones, through which pass myriads of electrical impulses, relayed across the synapses from one neurone to another by chemical transmitters or in some cases electrically. There are dozens of chemical substances that are now known to function as neurotransmitters, or as modifiers of their actions, including the nervous system's own internal pain-killers, the enkephalins and endorphins. The principles of the "wiring" of the nervous system are the same throughout the cortex, regardless of the functional specialization of a particular area. Nature seems to use the same basic building blocks over and over again, and employs no cells, circuitry or neurotransmitters unique to human beings. Organization determines function, and the modular principle extends downwards from regions of the brain that mediate

major functions to the cortical columns of neuronal structures. Changeux describes these discoveries with clarity and panache, making excursions into the brain chemistry of pain, thirst, rage and even orgasm.

Behaviour can be explained in terms of the mobilization of sets of nerve cells; their responses can be explained in physicochemical terms; and these interactions can be explained at the molecular level. But what about the subjective life of the mind, such as our ability to imagine a friend's face, our emotional experiences and our sense of self-awareness? Changeux's thesis is that mental states are identical to physical states of the brain. Percepts, images, concepts and all such "mental objects" correspond to the activity, electrical and chemical, of assemblies of neurones, dispersed through separate regions of the brain in the case of abstract ideas. Changeux admits that a given mental object may be constructed from slightly different neuronal populations, which may even differ in detail within the same individual from one moment to another. But he does not deal adequately with how the same mental state could arise from such different physical states of the brain.

The answer to this problem is to be found in a different species of materialism that can be traced back to Kenneth Craik and Alan Turing. What mentality depends on is not a particular physical substrate, but the functional organization of the processes that it makes possible. There is still no need to invoke mystical properties in explaining the mind, but this approach can be informed by the theory of computability. The brain is clearly very unlike an ordinary digital computer in the details of its construction and operation; yet it is arguably a computational device that is causally connected to the external world.

If one thinks of the brain as a system containing many processors that carry out computations in a parallel and distributed way, one can make sense both of Gazzaniga's modular hypothesis and of the relation between mentality and neurones — just as different algorithms can compute the same function, so the same mental state can arise from different configurations of neurones provided that the functional organization of their processes is the same. It is therefore necessary to understand what the mind's various computational tasks are, how they might best be carried out, and how such procedures can be neurally embodied. It is a pity that the work of Craik, Turing and the late David Marr, who did so much to clarify these issues, lies outside the scope of these two books. The brain is the organ of the mind, but the dependence cuts both ways. If you want to understand the brain, you had better understand the mind. □

P.N. Johnson-Laird is Assistant Director of the Medical Research Council Applied Psychology Unit, 15 Chaucer Road, Cambridge CB 2EF, UK.

Entropy parameters and quantum jumps

Walter Gratzer

Mayonnaise and the Origin of Life: Thoughts of Minds and Molecules. By Harold J. Morowitz. *Charles Scribner's Sons*: 1985. Pp.256. \$15.95.

Natural Acts: A Sidelong View of Science and Nature. By David Quammen. *Lyons Books (Schocken)*: 1985. Pp.221. \$16.95.

THE first of Dr Morowitz's essays finds him on the Galapagos Islands, the third in an aeroplane approaching New Delhi and not long after we meet him pacing the deck of a cruise boat on Glacier Bay in Alaska. Where other orbiting professors sort their slides or touch up the odd grant application, Dr Morowitz, it seems, reaches for his pen and kicks his muse into action.

As a popularizer of science Morowitz is by no means in the Haldane class, but, at his best, as when he is ruminating on the quirks of scientists, dead and alive, and the assaults on science by the Great Unwashed, he is distinctly good company. He writes captivately on the eccentric palaeontologist, John Bell Hatcher, and on the unhappy Philip Gosse, a Victorian worthy, who as a Plymouth Brother, a zoologist and FRS, and friend of Darwin, was tormented by the geological and palaeontological evidence against Bishop Ussher's chronology and by such teasers as whether the occupants of the Garden of Eden possessed navels. (Such matters were far more satisfyingly resolved by the mediaeval schoolmen: what could be more pleasing, for example, than the logic of Origen, who held that since the sphere was the most perfect of all shapes by virtue of the invariant distance between its centre and all points on the surface, we would all on the Day of Judgement be transformed into spheres and roll into paradise.) Morowitz respects men such as Gosse for their tortured honesty of intellect. He deals sharply, by contrast, with those modern fleas on the body of science, the creationists and parapsychologists. One of his most entertaining pieces is a magisterial analysis of the thermodynamics of ESP.

I am less sure about a stertorous attempt to rehabilitate Teilhard de Chardin. But then again Morowitz comes up trumps on popular perceptions of entropy, which seems to have joined parameter, extrapolate and quantum-jump in the journalistic vocabulary. Here is a political pundit, quoted by Morowitz: "Entropy helps to explain why we have runaway inflation, soaring unemployment, bloated bureaucracies, a widely escalating energy crisis and worsening pollution". Poor Boltzmann (over whose remains in Vienna is hewn in stone the inscription

New In paperback

The Evolution of Insect Mating Systems. By Randy Thornhill and John Alcock. *Harvard University Press*. £16.95, \$19.95.

Animal Thinking. By Donald R. Griffin. *Harvard University Press*. £6.95, \$7.95. which was reviewed in *Nature* 313, 410.

The Mathematical Theory of Quantitative Genetics. By M.G. Bulmer. *Oxford University Press*. £12.50.

$S = k \log \Omega$) surely did not foresee this.

Morowitz offers some agreeable literary diversions and his thermodynamite's view of the passage in *Ulysses*, in which Leopold Bloom boils a kettle for his shaving water, deserves to be preserved. He is also rather good on dentists: he enjoys the "intellectual and aesthetic pleasure" of feeling a craftsman at work in his mouth. (Bernard Shaw took the same view: there was nothing more soothing, he wrote in one of his music criticisms, after a harrowing series of recitals, than to have one's teeth drilled by a finely-skilled hand.) There is much in this collection to entertain, sometimes to annoy and only occasionally to bore.

David Quammen disarmingly confesses to being a dilettante rather than a scientist, "a haunter of libraries and a snooper". He writes natural history in an unmistakably American style and is given to apostrophizing his reader in sporting metaphors. "Truly", he says, speaking of the larger kind of sea cucumber, "these guys are out in left field". This, I own, stretches my tenuous grasp of baseball. And what, for that matter, is this Hostess Twinkie, that a seal allegedly resembles in section?

The sea cucumber opens Quammen's collection, and its trials are considerable: the tiny parasitic *candiru* fish, for instance, enters and leaves the cucumber with impunity by way of the nether passage, lingering only to feed on its vitals. "You'd think", Quammen exclaims, "the little bugger was selling encyclopaedias". Well, these are not the lapidary periods of a Medawar or Gould; but we do learn how the resourceful cucumbers deal with the situation when they eventually weary of the tiresome intruder: they evert themselves and "blast their own gut and internal organs out through their own anus". Humans are denied this recourse, and, says Quammen, when the *candiru* enters the urethra it lodges there, its spines facing backwards, and the only alternative to death by blood poisoning is amputation. Beginning to feel faint, I passed on to the next story, which is about bats, and here was rewarded by an account of a scheme by the American military during the Second World War to subdue the Japanese by dropping bats, each fitted with a napalm bomb and a little parachute, on their cities.

Quammen's book is full of strange information about animals and their habitat, about Tycho Brahe's artificial nose, and other such oddities. His exuberance carries one, sometimes unwillingly, along. He has a penchant for the poetry of Donne, and an informed concern for wildlife and the environment that is altogether heartening; his book deserves every success. □

Walter Gratzer is in the Medical Research Council Cell Biophysics Unit, Kings College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

Behind the scenes in the fight on tropical diseases

Adetokunbo O. Lucas

From the Face of the Earth*. By June Goodfield. *André Deutsch*: 1985. Pp. 256. £10.95.

THE public exhibit an insatiable hunger for news about medicine and the medical sciences. In this book June Goodfield provides a five-course offering to feed the layman's appetite for "stories of medical achievement". As popular science, it translates into everyday language scientific information that is often obscured by jargon. The story telling is enlivened by taking the reader behind the scenes of scientific events — we read not only of dedication and the determination to succeed, but of personal disputes, divorce, dementia and other such human problems.

The first story about Kuru contains all the excitement of a strange disease, uniformly fatal among the sufferers, with a maximal impact on women and children and threatening to wipe out a whole community. This account highlights the role of those who finally solved the mystery. At the top of the list is Carleton Gajdusek, the deserved winner of a Nobel prize for identifying the cause of Kuru and for the discovery of slow viruses as causes of human disease. June Goodfield, however, notes but does not stress the point that control of Kuru in Papua New Guinea did not result directly from the outstanding biomedical and behavioural research which she so vividly describes. Rather, the decline in the incidence of Kuru apparently stemmed from action taken by some person(s) unnamed, who, in the 1950s, forced a cultural change in the ritual method of disposing of the dead. Who,

*In the United States *Quest for the Killers*, published by Birkhäuser price \$24.95.

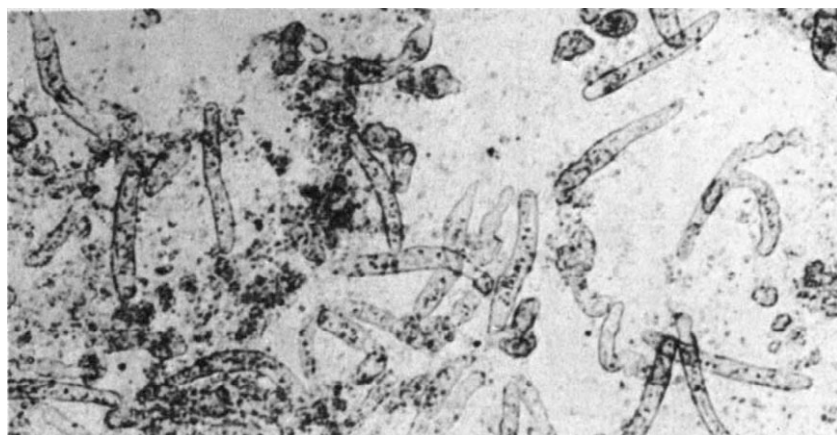
one wonders, were these faceless persons? What was their motive? Was it mere abhorrence of a strange cultural practice, or did they somehow intuitively sense that such practices could have fatal consequences?

The Kuru story illustrates the fall-out effect by which research on a specific problem can generate new knowledge of fundamental importance and unexpected benefit in other areas. Thus the work in Papua New Guinea has brought to Europe and North America new insights to diseases of local interest, such as Gerstmann-Sträussler Syndrome and Creutzfeldt-Jakob Disease.

Discipline, dedication and imaginative resourcefulness are the main ingredients of the next story on hepatitis B vaccine. Illustrated here is the point that the identification of the appropriate risk groups is an essential stock-in-trade of the epidemiologist, from classical studies on scurvy among English sailors, to research into cancer of the lung among smokers.

"The Three Valleys of St Lucia" deals with a large operational research programme aimed at comparing strategies for the control of schistosomiasis. Unlike Kuru, a unique disease in an island population, schistosomiasis is widespread in the tropics, affecting some 200 million people. It is not new — its eggs have been found in an Egyptian mummy — but it presents a fresh challenge where projects such as irrigation schemes and man-made lakes result in expansion of the population of the vector snails, an increase in man-water contacts, and thus intensification and broadening of the distribution of the disease.

The St Lucia story well illustrates the practical questions that need to be answered in translating theoretical options into effective strategies. And, again, the research provided answers that were useful not only on St Lucia but also in other areas. Recently, the World Health Organization proposed new strategies for the control of schistosomiasis which owe much to the findings at St Lucia, updated to take note of recent advances in technol-



Schistosome sporocysts in the head-foot region of a snail. From *Schistosomiasis: The St Lucia Project* by Peter Jordan. Cambridge University Press. To be reviewed in *Nature*.

ogy, in particular the availability of a new, well-tolerated drug.

The chapter on leprosy is interesting, but perhaps premature in comparison with the other stories. It tells of the search for a vaccine; but while much progress has been made over the past decade we are still far from finding a vaccine of proven efficacy for operational use.

The last story, on the fight against smallpox, provides a grand finale in that it fully justifies the title of the British edition of the book. This was a unique achievement, the total eradication of a disease that presented a worldwide challenge. Even though Kuru may disappear, one cannot sign its death certificate with the same confidence as was done for smallpox on 8 May 1980 at the World Health Assembly in Geneva. Although there has been some speculation about the feasibility of completely eradicating other infections "from the face of the Earth" — proposed candidates include measles and poliomyelitis — no other disease presents the combination of features which made possible the difficult task of eradicating smallpox.

Most of the data required for epidemiological surveillance of smallpox could be obtained from fairly simple clinical observations of sick patients, and the examination of scars of healed natural infections or vaccination scars. Laboratory help was required in the identification of the small proportion of atypical cases, and for the differentiation of smallpox from other pox diseases. Most other infections present more formidable problems; the clinical features are usually not sufficiently characteristic to permit differential diagnosis on simple clinical examination, and there are usually no pathognomonic scars to identify those who have had previous infections or are protected by immunization.

Apart from satisfying the demand for exciting accounts of medical discoveries, this book and similar ones draw the public into the debate about investment in the biomedical sciences. On the one hand, at times of economic recession, reduction in the budgets for biomedical research often seems a soft option because the venture does not command a powerful constituency. On the other hand, when threatened by the appearance of such problems as Legionnaire's disease and AIDS, people are anxious to know that something is being done to control the scourge. This book shows that breakthroughs do not leap out on the order of clever scientists. It portrays the path to knowledge as a difficult one, more often crossed by disappointments than by the discoveries that come as the final reward for long-sustained devotion to science and the improvement of human health. □

Adetokunbo O. Lucas is Director of the Special Programme for Research and Training in Tropical Diseases, World Health Organization, 1211 Geneva 27, Switzerland.

Lemons and limes

Joe Collier

Drug Discovery: The Evolution of Modern Medicines. By Walter Sneader. Wiley: 1985. Pp.392. Pbk £12.95, \$21.95.

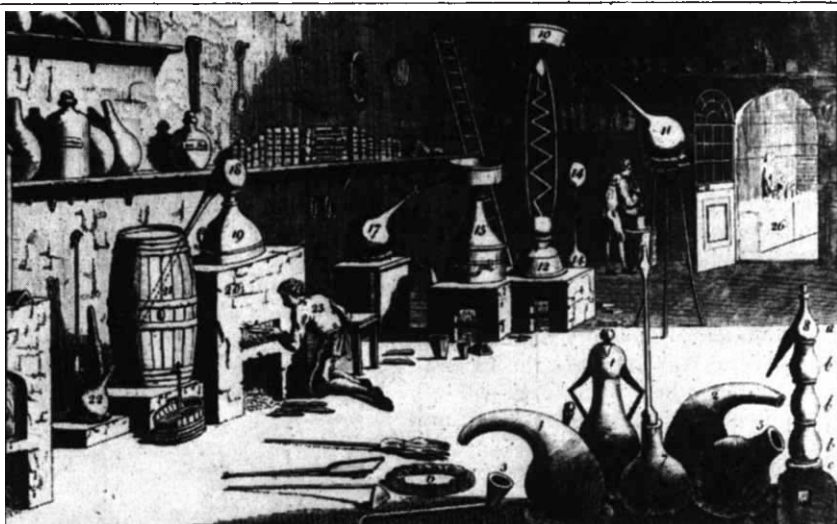
PHARMACOLOGISTS, physicians and legislators alike are now very wary of assuming that a drug is necessarily a medicine. Whereas a drug is any chemical substance that modifies a biological process, a medicine is that rare drug that can be used in patients to treat disease.

It is this distinction that occupies regulatory authorities (such as the Committee on Safety of Medicines, and the Food and Drug Administration) when licensing drugs for medical use; equally, it is the gap that the pharmaceutical industry strives to bridge so that new products can be marketed. Accordingly, as advances in fields such as biochemistry, genetics, immunology and pharmacology have led to the production of evermore sophisticated drugs, legislators have demanded increasingly detailed information about them before they will consider granting a licence. This balance now dominates the development of new therapeutic agents, and it is to this end that industry labours to discover how new drugs are absorbed, metabolized and excreted, how they should be formulated to be delivered to best advantage, how the dosage frequency should be adjusted to offer the greatest chance of efficacy, and, in longer term trials, the details of safety and efficacy.

Walter Sneader seems to be almost oblivious of these issues which are so central to the title of his book. They neither seem to be a consideration from which he might view the past, nor have they proved worthy of mention as a topic in their own right (and one so admirably dealt with by R.D. Mann in *Modern Drug Use: An Enquiry on Historical Principles*, published last year by MTP Press).

Instead, Sneader has filled the book with classical information. Each chapter (there are 16 in all) is devoted to a single broad therapeutic category, such as antibiotics, cardiovascular drugs, psychopharmacological agents, and in each the group is traced from the earliest reference to a seminal substance (penicillin, digoxin, Rauwolfia) through to drugs developed in recent history (the date seems to depend on Sneader's interests; inotropic agents stop with digoxin in 1933). The material is mostly in story form, in which the full names and addresses of each and every worker are listed, as are the movements of compounds and ideas between departments, companies and continents. There is an abundance of anecdotal information, clearly written and easy to follow — it is fascinating, for instance, to learn that scurvy returned as a threat to sailors when the trusty lemon juice was replaced by lime; lime juice, which was introduced for reasons of economy, has, we are told, about one-quarter of the ascorbic activity of lemon juice.

Two good reasons for publishing a new history book are that either it reveals fresh information or it analyses history from



Eighteenth-century manufactory of drugs, with a direct sales outlet to the public (far right). The illustration appears in *The Development of a Medicine*, by R.B. Smith, recently published by Macmillan. London (hbk £30; pbk £12.95)/Stockton, New York (hbk \$60).

Without fuss or pretension Smith's short book (160 pages) describes the development of medicines from their discovery (by serendipity or search), through their manufacture, their assessment both in the laboratory and the clinic, and on to the procedures involved in their marketing as licensed products. It gives a simple and straightforward account of contemporary practice in Britain, written by someone obviously familiar with the industry — when discussing drug company activity the book gives some rare glimpses of commercial philosophy in the pharmaceutical business. The book suffers from apparently being compiled in a hurry (tables illegible, legends wanting) but these should be corrected for the second edition that will surely come.

Joe Collier

novel point of view. Unfortunately, *Drug Discovery* does neither. The author's statements are always left open to question because no references are given in the text, so leaving the reader to search the rather poor and imprecise chapter bibliography at the end of the book. Interestingly, this sacrifice was made so that the text, which was written to convey "drama and excitement", should not be disrupted. However without references to contemporary information, the book has lost much scientific credibility.

Perhaps more annoying, there seems to be no benefit of hindsight. The ethical issues that have emerged so forcefully over the past 30 years and now engulf medicine seem to have left no impression. For its easy pace and the stories it relates I would expect the book to have its devotees; but as a serious historical account it falls short of what should be expected. □

Joe Collier is Senior Lecturer and Honorary Consultant in Clinical Pharmacology at St George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK.

A stir in the primeval soup

Jacques Ninio

Seven Clues to the Origin of Life: A Scientific Detective Story. By A.G. Cairns-Smith. Cambridge University Press: 1985. Pp.123. £9.95, \$17.95.

WE ALL admire the high technology feats that the simplest cells achieve in duplicating chromosomes, translating genes into proteins and adjusting their enzyme activities to their immediate needs. All such pathways are interdependent, each relying on all the others, as in an arch that would collapse if any one of its stones were missing. How did the first cells arise, if life requires, at a minimum, a translation apparatus to make proteins to replicate DNA?

The usual answer is that some 4×10^9 yr ago the Earth was a gigantic organic reactor producing masses of small molecules, amongst which were amino acids, sugars and nucleotides. Spontaneous polymerizations occurred, leading to nucleic acids that could replicate without the help of enzymes. Some of these "naked genes" acquired control over their environment. The coupling between nucleic acid replication and peptide synthesis became tighter and tighter until it became a strict coding relationship.

Cairns-Smith objects to this view; he believes that primitive organic reactors were more likely to produce tars and sludges than fine biochemical products, and that nucleotides are complex molecules that require sophisticated mechanisms for their production. Also, if a low-technology machine is to work, it must avoid the mutual dependency of its parts that is characteristic of more complicated machines. But, he says, the very existence of a primitive life-form made the synthesis of well-defined organic molecules possible. Amongst these, fancy polysaccharides, with no genetic use, might initially have been precursors of DNA or RNA. According to Cairns-Smith, new genes and the machineries they controlled gradually took over from the older forms of life: "None of the fibres in a rope has to stretch from one end to the other . . . new 'gene fibres' may be added and others subtracted without breaking the overall continuity".

These ideas are well known among scientists in general but are rarely discussed openly. In the origins-of-life research community, people seek a comfortable corner where they are least likely to interact with competitors. Yet, consider the fundamental issue: is nucleic acid replication conceivable in a prebiotic world? We are flooded with so-called theoretical treatments which generate models of the evolution of prebiotic self-replicating nuc-

Marking the path

Peter Laing

The Biotechnology Business: A Strategic Analysis. By Peter Daly. Frances Pinter, London/Rowman & Allanheld, Totowa, New Jersey: 1985. Pp.155. £16.50, \$25.

THE emergence of an international industry based on molecular biology and immunology is a striking example of the potential economic benefits of government funding of basic research. Equally, however, the commercialization of biotechnology well-illustrates the familiar pattern of technological innovation being avidly and rapidly exploited in the United States and only belatedly in Europe.

In the United States, the availability of venture capital and an entrepreneurial business climate have led to the formation of numerous specialist biotechnology companies. These have been able to attract many of the brightest academic scientists (who would not normally have considered an industrial career) and concentrate upon the problems of developing marketable products. The large, established American companies, aware of the many studies that have shown that technical innovation most often stems from small, flexible and highly-specialized firms, have generally encouraged this process by the provision of research contracts (and even equity finance) for start-up companies, obtaining in exchange the commercial rights to future products.

By contrast, in Europe the start-up company as an engine of technological change has until very recently been a rarity. The brunt of commercialization of new technology has been borne largely by the established industrial companies, in which new ideas have a depressing tendency to be stifled or restrained by corporate inertia and the "Not Invented Here" syndrome. This situation has been compounded by a general lack of contact between industry and academia, particularly in the biological sciences, with the result that the more enterprising scientists have either remained in universities or left for North America.

In Japan, a strong sense of national purpose and a cultural preference for con-

sensus solutions have engendered active cooperation between Government, industry and academia with the aim of establishing national priorities for the exploitation of new ideas and identifying the means of bringing them about. Even so, the vehicle for commercialization in Japan has invariably been the established company — indeed, companies in moribund or declining industries have often been selected for revitalization by the injection of funds for biotechnology projects.

The Biotechnology Business is a comprehensive and readable guidebook to the commercial side of the biotechnology industry. As acknowledged by the author, it leans heavily on the excellent Office of Technology Assessment report *Commercial Biotechnology*, published in 1984, but is less nationalistic in its approach (and also weighs only a quarter of the OTA report's daunting 1.2kg).

Daly's emphasis on strategic analysis, made explicit in his sub-title, has resulted in many interesting case studies. Among them is a timely comment on the American company Genex, warning of the dangers of the company's heavy reliance on a single manufacturing contract for L-phenylalanine with G.D. Searle. This contract was terminated a few weeks after the book was published, plunging Genex into a financial crisis.

The source material for the book appears to have been culled largely from articles in specialist biotechnology magazines. These have tended to concentrate on the newsworthy start-up companies, and there is in consequence an under-representation of the strategies which are being adopted by the large chemical and pharmaceutical firms; now that the basic technology is more widely disseminated, such companies are beginning to make real inroads into the market. The likely impact of these moves on the start-up companies suggests not only a volatile employment market for the scientists involved, but, as Peter Daly points out, the development of an even more fruitful area for strategic analysis than the biotechnology industry has provided to date. □

Peter Laing is an Assistant Director of N.M. Rothschild Asset Management Ltd, New Court, St Swithin's Lane, London EC4P 4DU, UK, which acts as investment adviser to Biotechnology Investments Ltd.

leic acid populations as though the very existence of such reactions were not the primary theoretical issue. With highly-skilled organic syntheses, the beginnings of something like non-enzymatic DNA or RNA replication are observed. But now, as Schwartz and Orgel have said (*Science* 228, 585–587; 1985), we need to “design reactions that proceed efficiently with more plausibly prebiotic substrate mixtures”. Sooner or later, organic polymers, able to replicate in the laboratory and having simpler backbones than DNA or RNA, will be synthesized. Then the idea of “takeover”, as Cairns-Smith has called it, will become inevitable.

Cairns-Smith, however, is far more radical than this, in that he rejects the concept of any primitive *organic* gene. Instead, he thinks in terms of clay particles, made of successive layers, each new layer added on top reproducing the defect structure of the previous layer; or of clays made of a succession of various kinds of uniform layers, that grow sideways. If such particles break across the layers, then the one-dimensional information contained in the particular succession of layers is conserved. Ways in which clays might influence organic reactions and use them for their growth are discussed in this book. Thus clay particles are suitable substrates for natural selection. Will field studies be conducted on the population structure of clay particles in given ecological niches? Or will some courageous young scientist design a clay-chemostat where the growth of clay crystals will be studied under various selective pressures?

Cairns-Smith's ideas deserve to be

assimilated and discussed. In *Seven Clues to the Origins of Life*, a layman's version of his earlier book *Genetic Takeover*, all his essential theses are presented in a clear, precise and condensed manner. The choice of a deductive form of presentation for his arguments (clays are introduced only three chapters from the end), buttressed by references to Sherlock Holmes, lends much rigour to the plot, and leaves no escape exit. How can low-technology life be designed, and how can it evolve towards the high-technology form? I know of no other book that succeeds as well as this one in maintaining this central question in focus throughout. It is a summary of the best evolutionary thinking as applied to the origins of life in which the important issues are addressed pertinently, economically and with a happy recourse to creative analogies.

The concept of “takeover” is an important evolutionary one, and may apply to a wide variety of situations. I regret in this respect that the author did not attempt to analyse concrete biological examples (such as the origin of split genes, RNA catalysts and ribosomes, or the design of the immune system) in the light of this idea. Inevitably, I think, schemes in which things acquire complexity of detail without a change in their organization will have to compete with “takeover” schemes. The time has now come for those scientists working on the subject of the origin of life to cease dealing with such ideas obliquely. □

Jacques Ninio is at the Institut Jacques Monod, Tour 43, 2 Place Jussieu, 75251 Paris Cedex 05, France.

that may occur in any population of individuals”. He transferred this anti-essentialism from entomology to human behaviour. Human sexual behaviour cannot be pigeon-holed; just as the world of wasps was a continuum in each and every one of its aspects, so too the male population cannot be divided into two discrete groups, heterosexual and homosexual.



The enigmatic smile of a swan — or an upside-down flamingo?

Gould teases out the wider implications of a particular scientific enquiry, from the taxonomy of wasps to a more accurate — and tolerant — understanding of what is regarded as sexual deviation. The scientific, moral and political (Kinsey had the misfortune to coincide with Senator McCarthy) implications are shown to interact.

The essays range widely from speculation about the reason for the extinction of the dinosaurs to the bizarre and unhappy cases of the Hottentot Venus and of Ritta-Christina, the celebrated siamese twin sisters. The style is relaxed but not flip; the content all the better for being disciplined by the limits of the monthly column. If there is a weakness it is Gould's rather too obvious wish to hook his reader with a startling opening paragraph. An essay on continuity and quirkiness in evolution leads off with Michelangelo and one on the flamingo's eating habits with Buffalo Bill. It is good journalism but unnecessary: in all these essays the subject matter is compelling enough in itself.

Scientists who communicate their enthusiasm to a wider audience are regarded by some of their professional colleagues as vulgar; if it can be explained to a layman it cannot be serious. Gould brushes aside that mixture of arrogance and insecurity. His essays popularize science in the same sense that Macaulay popularized history — by making it interesting and entertaining to read. Roll on volume five. Meanwhile I shall be looking for copies of volumes one to three. □

John Rae, 17 Dean's Yard, London SW1P 3BP, UK, is Head Master of Westminster School.

Appeal to the laity

John Rae

The Flamingo's Smile: Reflections in Natural History. By Stephen Jay Gould. W.W. Norton: 1985. Pp. 476. \$17.95. To be published in Britain in February 1986, £12.95.

It is 20 years since C.P. Snow described the gulf fixed between scientists and non-scientists in academia. Not much has changed. The two communities continue to go about their business like a couple of Englishmen travelling in a foreign land: they are on the same journey but have nothing to say to one another because they have never been introduced. The non-scientist complains that the scientist cannot communicate what he is doing in a language that the layman can understand. The scientist complains that the non-scientist defines language in narrowly non-scientific terms — why should a biologist have to explain what he means by morphology any more than a grammarian has to explain what he means by tautology? No doubt it is the education systems

that are to blame; meanwhile the dialogue of the deaf continues.

Stephen Jay Gould overcomes this problem, not by laicizing the language, but by using it in the context of intellectual enquiries that any intelligent person can understand and by a technique of cross-referencing that relates the scientific enquiry to comparable problems in other disciplines. It might be argued that as a palaeontologist and evolutionary biologist he has an easier task than say, the nuclear physicist. But easy or not he succeeds brilliantly. This collection of essays, the fourth to be based on his monthly column in the *Natural History Magazine*, communicates both the fascination and the universality of scientific study in a way that can hardly fail to intrigue scientist and non-scientist alike.

Take, for example, Kinsey and the wasps. Kinsey's reports on sexual behaviour were not the product of his desire to promote sexual enlightenment. He drifted into sex research by accident and applied to the question his experience as America's leading taxonomist of wasps. His study of wasps persuaded him — in his own words — “of the uniqueness of individuals and of the wide range of variation

Biology and the behaviour of man

John Maynard Smith

Vaulting Ambition: Sociobiology and the Quest for Human Nature.

By Philip Kitcher.

MIT Press: 1985. Pp. 447. \$24.95, £24.95.

DO WE really need another critique of sociobiology? In general, probably not, but perhaps we need this one. Kitcher, like everyone else, approaches the problem with prejudices, but he tries harder and more successfully than most to rise above them. Prejudices are inevitable. It is natural for geneticists and evolutionary biologists to hope that their disciplines will throw new light on the human condition, and equally natural for social scientists to resist the threatened takeover. More important for many of us, previous efforts to apply biology to human affairs have too often ended up as justifications for racial, sexual and class inequalities. Kitcher, who grew up in England, has not forgotten that, in the post-War years, schoolchildren were divided at the age of eleven into sheep and goats, and that this division was justified by the leading experimental psychologists of the day. He and I share this experience — he as a tested child and I as a parent of tested children. It has left us cautious about proposals to use biological theory to plan human institutions.

Kitcher, then, is unsympathetic to the claim that evolutionary biology can guide

... what we now need in sociobiology is a more cautious analysis of the data ...

political judgement, and I suspect he was unsympathetic before he started work on this book. Unlike some other authors, however, he has undertaken a genuine study. He does understand the ideas he is criticizing. He has the biological knowledge to evaluate the evolutionary background to sociobiology, and the mathematical ability to analyse the claims made for it. Above all, he presents sociobiology in its strongest and most coherent form, and avoids the easy option of attacking only its more idiotic manifestations.

He distinguishes sharply between the attempt to understand the evolution of social behaviour in animals, and attempts to understand man. He is sympathetic to the former enterprise. Correctly, he points out that there is no special underlying theory: "There is no autonomous theory of the evolution of behaviour. There is only the general theory of evolution". It may be that interactions between relatives, and frequency-dependent fitnesses, were more important in the evolution of the behaviour of birds than in the evolution of their wings, but they are not peculiar to behavioural evolution: kin selection and game theory are just as relevant to plant evolution.

There is, of course, good and bad work in animal sociobiology, and Kitcher gives examples of both. The bad, he points out, has two characteristics: data are quoted as supporting some specific hypothesis, without considering alternatives, and the hypotheses themselves are modified after the fact until data and predictions are brought into line. However, his chapter "Dr Pangloss's Last Hurrah", which takes issue with the "adaptationist program", seems to me only partly correct. He presents two genetical reasons for not expecting perfect adaptation. The first is that there are genetic systems, even with constant fitnesses, in which selection will not fix the fittest genotype. The simplest is that of heterozygous advantage: if Aa is fitter than AA or aa, selection cannot produce a population consisting entirely of Aa individuals. This is of course true, but is it interesting? If we want to understand why some species does not have the phenotype predicted by theory, this kind of genetic detail is rather unlikely to be the reason. Suppose, for example, we are interested in the shape of vertebrate wings. Aerodynamic theory shows that the optimal shape is usually elliptical. Pterodactyls, however, never had elliptical wings, but no one would explain this by suggesting that, perhaps, only heterozygotes had elliptical wings. The true explanation has to do with the way in which pterodactyl wings were made. Of course, this would be reflected in the absence of certain kinds of heritable variability, but that is not a useful way of thinking about the problem. For phenotypes of the complexity typically discussed by sociobiologists, it is usually better to think at the level of development and physiology than of genetics.

All the same, there will be cases in which this kind of genetic constraint will be relevant. The second kind of genetic constraint discussed by Kitcher, however, seems to me to be a misunderstanding. He points out that when fitnesses are frequency-dependent, the mean fitness of a population may decrease under selection. Therefore, he says, "proponents of optimization analyses who show that a certain design would maximize mean fitness may not automatically assume that selection can produce this design". Now, as Kitcher understands very well, a major thrust of sociobiology has been to show that selection acting at the level of the individual does not necessarily lead to the evolution of characteristics optimal for the population. One of the main reasons for

this is that fitnesses are frequency-dependent. Evolutionary game theory was developed specifically to analyse such cases. It is not sociobiologists who suppose that selection "maximizes [population] mean fitness". It is ironic that the phrase "Pangloss's theorem" was first used in the debate about evolution (in print, I think, by myself, but borrowed from a remark of Haldane's), not as a criticism of adaptive explanations, but specifically as a criticism of "group-selectionist", mean-fitness-maximizing arguments.

Thus I think that Kitcher is unfair to sociobiologists when he introduces the argument from frequency-dependence. However, it is a rare slip: in general his

... It is natural for geneticists and evolutionary biologists to hope that their disciplines will throw new light on the human condition ...

account is just. I agree that what we now need in sociobiology is a more cautious analysis of data, and a more careful consideration of alternative hypotheses. This will not come easily, for reasons that are as much sociological as scientific. The critique by Gould and Lewontin has had little impact on practitioners, perhaps because they were seen as hostile to the whole enterprise, and not merely to careless practise of it. Theorists like myself are understandably delighted when some set of observations seems to fit their theories. Field workers, equally understandably, are pleased if their data receive a rational explanation. The time has come, however, for editors and referees to place more emphasis on the quality of the data, and the care with which alternative explanations have been considered, and less on success in fitting the data to some particular theory. But, as Kitcher insists, the enterprise is worthwhile, and the best work is of a high standard.

What of man? Clearly, biology must have something to say. Man is an animal, and has evolved by the same processes as other animals. The debate is between those who, while accepting that man is an animal, argue that he is such a peculiar animal that evolutionary biology can have little to say about his social behaviour, and those who think that the study of human societies, just as of ant societies, must be rooted in biology. The second position Kitcher refers to as "pop sociobiology". I think this is a pity, for two reasons. First, it gives an image of superficiality and appeal to popular prejudice which, at least sometimes, is quite unfair: it is hard to imagine anyone less "pop" than Richard Alexander. Second, it gives the wrong impression of what Kitcher himself is doing: he is scrupulous about putting the best interpretation on sociobiological arguments. But I see his difficulty: we do need a term for the application of sociobiology to human beings, and I have no better one to offer.

Kitcher's basic position is that one cannot dismiss pop sociobiology simply by asserting that it assumes genetic determinism and is therefore false, since plausible sociobiological arguments can be developed which do not assume that genes determine behaviour. There is therefore no escape from considering these arguments in detail, and to see if they stand up

... one cannot dismiss pop sociobiology simply by asserting that it assumes genetic determination and is therefore false ...

and deliver any fruit. Kitcher distinguishes two schools of pop sociobiology, those of E. O. Wilson and of Alexander. Wilson's basic argument he sees in the form of a ladder, as follows:

(i) We can plausibly argue that the members of some population, G, would maximize their fitness by exhibiting behaviour B.

(ii) If we observe that members of G in fact do B, we conclude that B became, and remains, prevalent through natural selection.

(iii) Because selection is effective only if there are genetic differences, we can conclude that there are genetic differences between current members of G and their ancestors, who did not do B.

(iv) Because there are genetic differences, and because the behaviour is adaptive, the behaviour will be difficult to modify by altering the social environment.

This is a shortened version of Kitcher's reconstruction. A major part of his book consists of a step-by-step critique. Clearly, the last step is based upon the shakiest grounds: the fact that our ancestors did B in all previous environments is not proof that they will do B in a wholly new one. The first three steps look more secure. Kitcher's most effective criticism here is not of the logical possibility of taking these steps, but of the ways in which they are in fact taken. For example, consider sexual behaviour. He quotes Wilson as espousing the view that evolution will lead to males that are "aggressive, hasty, fickle, and indiscriminating", and females that are "coy". But theory suggests not one evolutionary optimum, but several, and a number of our primate relatives form long-term pair bonds and show extensive male parental care. Hence there is little justification for Wilson's first step onto the ladder.

A second illegitimate way of getting on to the ladder is to apply to animals words which describe some human behaviour. For example, mallard drakes are said to "rape" ducks. Now it is true that drakes do force copulations on ducks and, by so doing, probably increase their fitness. What is the harm in calling this rape? If you are interested in ducks, rather little, but if you are interested in people, quite a lot. It implies that human rape occurs because it increases the inclusive fitness of the rapist. The contexts in which rape

occurs makes this implausible. I agree that there is a danger in applying words such as "aggression", "incest", "homosexuality" and so on to animals and man alike, when the behaviours referred to may be quite different. However I do have reservations. The alternative is often to invent a turgid and incomprehensible vocabulary to describe what animals do. I remember our unsuccessful attempt to introduce the term "kleptogamy" at an ethological congress, because we feared that the Anglo-Saxon alternative might offend our hosts.

A few years ago, I worked through the equations in Lumsden and Wilson's *Genes, Mind, and Culture* and found them to be badly flawed. Kitcher has done a still more thorough job, and come to essentially the same conclusion: his chapter is entitled "The Emperor's New Equations". On three occasions Wilson has found it helpful to find a mathematical collaborator. His first two were Robert MacArthur and George Oster: he was third time unlucky.

The second approach to human sociobiology, taken, for example, by Alexander, Irons, Chagnon and Dickemann, is more direct. Man is treated like any other animal. The question asked is as follows: given the social environment, do people behave so as to maximize their inclusive fitness? The answer, it is claimed, is "yes". Unlike Wilson's arguments, which seem to me generally ill-formulated and empty of content, this claim is worth taking seriously, even though it is probably false. Kitcher, attacks it on two fronts. First, he asks what proximate mechanism could possibly bring such behaviour about, since it seems to require an unconscious relationship-calculator and fitness-maximizer influencing our conscious actions. If I were Alexander, I would reply that, if the claim is true, then it is up to psychologists to discover the mechanism.

Kitcher's second line of attack is to ask whether people do in fact maximize their fitness. Here, the test case is Dickemann's account of societies practising female infanticide, and other acts not obviously contributing to fitness. Kitcher gives a careful analysis of this case, and develops a mathematical model of it suggesting that the increases in inclusive fitness that Dickemann proposes would not in fact occur. I am not sure whether he is right, but this is where the action is. This school of sociobiologists do say things about real societies that are testable; I find it hard to believe that they are right, but at least they are not vacuous.

This is an admirable book. Kitcher has the necessary background in biology, mathematics and philosophy. He is aware of his prejudices, and does his best to overcome them. This will not be the last word, but it is the best one yet. □

John Maynard Smith is Professor in the School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, UK.

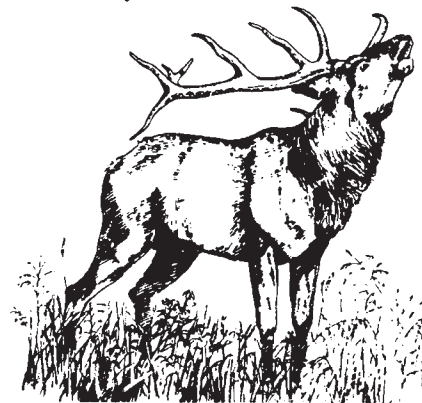
A case of peripatetic mammaldom

David W. Macdonald

Naturalized Mammals of the World. By Christopher Lever. Longman: 1985. Pp. 487. £40. To be published in the United States late 1985/early 1986, \$79.95.

THE innocent naturalist might relish a glimpse of an elusive mammal as a glimpse of unspoiled nature — a sight that might have brought a thrill in just that place, under just those circumstances, for millennia. Surely mammals, of all creatures, with their unassailable wildness, and secretive and often nocturnal habits, remain untainted by the meddling hand of man, albeit restricted to dwindling enclaves of suitable habitat?

Naturalized Mammals of the World shatters any such illusion that wild mam-



Red deer — now a New World resident.

mals are where they ought to be, or rather where they used to be. The reader will not be many pages into this intriguing catalogue of mishaps before succumbing to the overwhelming impression of a peripatetic mammaldom, of which an astounding array of displaced species are wreaking biogeographical havoc. Most people could name a handful of aliens in their country — in Britain, for example, the grey squirrel and coypu — but common knowledge of the few belies the many less-notorious species which have been deliberately (and, generally, in retrospect, misguidedly) shuffled around the globe. I got up to 67 major introductions before losing count. To qualify for the doubtful distinction of inclusion in the book, a species should have been imported from its natural range to a new country or region either deliberately or accidentally by human agency, and should currently be established in the wild in self-maintaining and self-perpetuating populations unsupported by, and independent of, man.

Apart from a succinct introductory section, and a set of useful tables and appendices, the book is made up of single-species monographs, each of which documents that species' displacement and sub-

sequent misdeeds. An addictive feature of these monographs is the pair of contrasting world maps provided, showing the animal's range "before" and "after" the onset of globe-trotting; the temptation, which I suspect few will resist, is to thumb ahead for a furtive peep to discover whether your favourite species suffers the stigma of far-flung black blobs scattered improperly across the "after" maps. The text itself is an enjoyable blend of biology and history, and is written in a precise and scholarly style. Although the author does not dwell on the implications for ecological theory of the effects of different immigrants on their new habitats, his text draws attention to the extraordinary interest of the process of invasion, and the many "unnatural" experiments that have been inadvertently established worldwide.

The history of three canids illustrates the scope of the book and the variation in the examples it documents. Red foxes were introduced to Australia in the late nineteenth century, in numbers generally no greater than two's and three's, by would-be hunters. Their progress can be charted in detail year by year from local records, until they are now widespread to

the extent that tens of thousands of descendants of the original few are killed annually. (Incidentally, British foxes were also shipped across to populate the eastern United States, and foxes from Spain were imported to bolster the numbers in Britain.) Then there is the raccoon dog, which originated in the Far East. For them, the scale of the operation was quite different. Between 1929 and 1967, 8,651 were released in European Russia by those with an eye on their pelts — in 1947 alone, 1,147 gravid females were let loose in the Caucasus. Each of these cases differs from that of the dingo, which was probably shipped to Australia some 3,000 years ago and so must soon qualify for honorary citizenship.

Although essentially a reference volume, this book has the "Guinness Book of Records" type of appeal to a browser. Apart from serving its main purpose, it will doubtless provide an abundance of worthwhile "Did you know . . . ?" gambits for both amateur and professional naturalists. □

David W. Macdonald is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Botanical spread

Brian Mathew

A Guide to the Vegetation of Britain and Europe. By Oleg Polunin and Martin Walters. *Oxford University Press*: 1985. Pp.284. £17.50, \$25.

AMID the dense rash of wild-flower books dealing with the flora of Britain and Europe at the species level, it is refreshing to be presented with a volume devoted to whole communities of plants. Polunin and Walters are probably accurate in claiming that this is the first attempt to bring together all the information on the subject at a layman's level, although, as they acknowledge, there are many other detailed works covering individual countries or regions.

Accordingly, a helpful bibliography has been included, divided into vegetation types, for those who wish to pursue any interesting points raised in the text. For internal reference, the reader is also well provided for, with indexes to Latin names, English names and plant communities, and a useful glossary (although I was surprised to find no mention, here or in the index, to the commonly used term "downland"). Conservationists, who now tend to put as much emphasis upon the preserva-

tion of habitats as of individual species, will find much useful material is provided, and there is a descriptive list of the Nature Reserves and National Parks throughout Europe.

The opening chapters deal with the all-important subjects of soils and climates, which have a great bearing on the plant cover, and there is an interesting summary of the history of Europe's vegetation from the earliest fossil records onwards. The bulk of the book consists of 14 chapters devoted to different regional zones — Arctic, Mediterranean, Central Europe and so on — and, within these, there are subdivisions down to communities with examples of the species to be found in them. The text is supported by many good line drawings and a few rather poorly reproduced monochrome photographs; the latter, however, are more than compensated for by the 110 excellent colour plates of vegetation types.

The authors have faced a difficult task, and it is not surprising that the book may appear bitty at first sight. As with most aspects of the botanical world, the demarcation lines are often ill-defined and there is much overlap, for example between *maquis* and *garigue*, with several different types of each of these. To have presented an over-classified statement of the situation would have led to many inaccuracies, and it is to the particular credit of Walters and Polunin, given the intended audience of their book, that they have struck the right balance between the superficial and the overly learned. □

Brian Mathew is a Principal Scientific Officer in the Herbarium, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK.

Animals and plants first

Michael J. Bean

International Wildlife Law. By Simon Lyster. *Grotius, Llandyssul, Dyfed SA44 4BQ, UK*: 1985. Pp.470. Hbk £25, \$37; pbk £12, \$17.50.

THE earliest efforts at international cooperation in the conservation of living resources date back well into the nineteenth century. Since then, the number of international treaties and agreements aimed at conserving wildlife has grown steadily. Not until Simon Lyster's book, however, has anyone tried to offer a comprehensive analysis of that body of law. The result is a compelling display of legal scholarship that should reward scientists interested in policy matters and conservationists around the world.

This is a book by a British lawyer about international conservation law. It escapes, however, the trap of appealing only to a narrow audience of similar specialists. After an introductory chapter that provides a layman's overview of certain general principles and characteristics of international law, Lyster turns his attention to a series of wildlife treaties. For each of these he offers some historical background, including a discussion of the problems that the treaty was intended to address, a detailed analysis of the terms of the treaty, and an assessment of the treaty's actual implementation and effects.

The most remarkable aspect of the book is its sheer scope. It is not limited to the treaties of a particular country or geographical region, and reaches out across ocean barriers to encompass the Convention on Nature Protection and Wildlife Preservation in the Western Hemisphere (1940), the African Convention on the Conservation of Nature and Natural Resources (1968), a pact of 1969 among Andean nations for the conservation of the vicuna, an agreement of 1973 among circumpolar nations for the conservation of the polar bear, numerous European treaties for the conservation of birds and other wildlife, and yet other initiatives.

The most thorough analysis in the book is devoted to four treaties of the 1970s that Lyster calls "the centrepiece of international wildlife law": the Convention on Wetlands of International Importance Especially as Waterfowl Habitat (1971); the Convention Concerning the Protection of the World Cultural and Natural Heritage (1972); the Convention on International Trade in Endangered Species of Wild Fauna and Flora (1973); and the Convention on the Conservation of Migratory Species of Wild Animals (1979). The complete texts of these and eight other treaties are reprinted in full in

● *Flowering Plants of the World*, consultant editor V. H. Heywood, which was published in 1978 by Oxford University Press and has been out of print, has been re-issued by Croom Helm in Britain and Prentice-Hall in the United States. Price is £19.95, \$39.95.

an appendix that occupies nearly a third of the entire volume.

The criteria upon which Lyster bases his judgement that the four treaties he principally addresses are in fact the most important are unstated. All are indeed global, open for signature by any nation of the world, but only two, the World Heritage Convention and CITES (the Endangered Species Convention), have the adherence of a majority of the world's nations. The Migratory Species Convention may be the most recent of international conservation initiatives, but with only 15 party states and no real accomplishments to date, Lyster's claim that it occupies part of the "centrepiece of international wildlife law" is premature at best. The Wetlands Convention, though it has served to direct attention to a particularly pressing problem, is notable mostly for its failure to create clear, enforceable obligations. While it is dealt with elsewhere in the book, the International Convention for the Regulation of Whaling might more properly have been included in the centrepiece if world attention and actual results were the criteria for that uncertain distinction.

A recurring theme throughout the book is the wide, sometimes enormous, gap that separates the aims and apparently literal meaning of most of these treaties and their actual implementation. For example, the World Heritage Convention tries to come to grips with the reality that many of the world's great cultural and natural resources are in the countries least able financially to assure their protection or forego the short-term benefits associated with their exploitation. The mechanism fashioned to answer this dilemma is, through the creation of a trust fund, to transfer wealth from the developed to the developing world for purposes of conservation. Though there is nearly universal agreement that such a transfer is needed if the intentions of the Convention are to be realized, the actual amounts of money made available for the purpose have been so small as to undermine any confidence that the commitment to the principle is anything more than cosmetic. Beyond noting the fact of such gaps between promise and performance, Lyster does not delve into explanations or suggest possible solutions.

His book stands, however, as an impressive survey of some of the efforts at international cooperation made over the past century to stem the loss of wildlife. It is, at the same time, a depressing reminder of how short many of those efforts have fallen. □

Michael J. Bean is Chairman of the Wildlife Program of the Environmental Defense Fund, 1616 P Street, NW, Washington, DC 20036, USA.

• Reviews of *Mind and the New Physics* by Fred Alan Wolf, *the Dialectical Biologist* by Richard Levins and Richard Lewontin, and *The Background of Ecology* by Robert P. McIntosh will appear in future issues of *Nature*.

More Darwinian detractors

Mark Ridley

Evolution: A Theory in Crisis. By Michael Denton. Burnett, London: 1985. Pp.368. £17.50. To be published in the United States in April 1986 by Adler & Adler, Washington, DC.

Creation and Evolution: The Facts and Fallacies. By Alan Hayward. Triangle, London: 1985. Pp.232. Pbk £2.75.

Adam and Evolution. By Michael Pitman. Rider, London: 1985. Pp.268. £9.95.

Evolutionary Theory: The Unfinished Synthesis. By Robert G.B. Reid. Croom Helm/Cornell University Press: 1985. Pp.405. £25, \$34.50.

WRITING a review of some long since forgotten anti-Darwinian work in this journal 95 years ago, Raphael Meldola remarked that "it is hardly necessary to say that many of the most weighty objections have been culled from the writings of Darwin himself". *Toutes ces la même chose*. Another four books published during 1985 now confirm that Mr Darwin's critics are as active, and their procedures the same, as ever. I do not think it is worth replying to them here. Readers of *Nature* can scarcely need to be put right on these old issues; and, anyhow, Darwin himself and many others have dealt with them at the length they deserve. How can I hope to succeed with three authors (Denton, Hayward and Pitman) who, like the Victorian astronomer Sir John Herschel, think that evolution by natural selection is the "law of higgledy-piggledy" — a "random search mechanism" (Denton), of "pure chance" (Hayward and Pitman)? Or with another author (Reid) who, like the Duke of Argyll, thinks that there must be some grave defect in Darwin's theory, because it personifies Nature in its analogy with artificial selection?

Denton, Hayward, Pitman and Reid, then, are opposed to the Darwinian theory. The first three of them (as we shall see) object to it for similar reasons; but they differ in what they would put in its place. Hayward is a creationist and Christian. Pitman is a creationist too, and probably a Christian one, although he has heretical leanings towards the dualistic (not trinitarian) wisdom of the East and an enthusiasm for music mysticism. Denton's purpose is purely destructive: he has no alternative to offer. Reid thinks he has a secular alternative, which we shall come to. His is the only book of the four that is written in a cerebral style, for an audience of professionals of some kind. Denton, Hayward and Pitman all write for uninformed readers, the latter two in the kind of short, didactic sentences and superficially patient tone that, in Britain, one associates with the closed minds of their

school. Denton has a different style. His pen is constantly running out of control. Every few pages he treats us to some more or less silly exaggeration. The preface sets the tone by telling us that "every aspect of evolution theory is being debated with intensity", at conferences, in journals and in the Natural History Museum in London. Denton may think that exaggeration is a necessary part of popular writing. If so, he is wrong.

None of the authors is an evolutionary biologist. Denton is a biochemist; Reid a philosophical physiologist; Hayward a retired physicist; Pitman "has an MA in classics". They would not think this a disadvantage. A Darwinian education (they believe) cripples the mind, filling it with emotional prejudices. Hayward accordingly prefers the unbiased view of the "outside observer", and Bernard Stonehouse without irony recommends Pitman to us as follows: "his unorthodox background must have helped Michael Pitman to write this book; it might not have occurred to one formed in a more conventional mould".

Their procedure is to sift through the writings of Darwin, and such popular secondary and tertiary sources as Stephen Gould's essays, *New Scientist* and even *The Guardian*. From this material, they seize upon the bits that look like difficulties for Darwinism, and ignore everything else. Then, after surrounding the difficulties with schoolroom rhetoric, sub-Kuhnian psychobabble and suitably simplified Victorian history, they send the whole to press. As argument, indeed, all four books are sad stuff. The reasoning I shall come to, but mark well the style. Denton is especially keen on the device (it must have a technical name) "even Darwin-Simpson-Sir Peter Medawar admits—concedes—acknowledges that . . .", which he places before each attenuated quotation, in order to reverse their meanings. I similarly found myself, in Hayward, "grudgingly admitting" something that in reality I unhesitatingly espouse. Denton's characterization of Lyell provides another example. "For most of his life", we are told — and this is all we are told — "[he] was vigorously opposed to the idea of evolution".

So why do Denton, Hayward and Pitman object to the neo-Darwinian theory? Their case (the arguments of the three are similar) is essentially Victorian. It plays on two main "difficulties": "chance" and "gaps". Natural selection, they all allow, can cause small changes in simple adaptations; but they deny that it can cause large changes among types and the evolution of complex adaptations like the vertebrate eye. "Different processes altogether may be involved, and here Darwin's strength fails" (Pitman). As they have identified natural selection with chance, they deduce (correctly) that it could not in principle produce large directed changes. Only the premise is wrong. In the Darwinian

theory, natural selection is not a chance process and complex adaptations evolve in many small changes. That possibility they either silently ignore, or misunderstand and quickly brush aside.

Then we have the problem of gaps. There are gaps between the classes of modern forms, and those well-publicized gaps in the fossil record. The origin of life is an example. According to these authors, it would not be possible for evolution by natural selection to cross the gap between the non-living and the living. After all, where are the intermediate forms? They have not been found, even though, according to Pitman, "billions of dollars are spent annually in attempts to demonstrate abiogenesis under laboratory conditions". Pitman also quotes Francis Crick, who "admits" that "regarding the origin of life we often find 'too much speculation running after too few facts'"; but speculative pursuit is as ill-advised when the aim is to demonstrate a negative, as a positive, conclusion. Darwin's own remarks about difficult transitional stages — that they are a problem of the imagination, and not of the reason — is silently ignored, but implicitly illustrated.

The theory of "punctuated equilibrium" of Eldredge and Gould here again does time for the creationist cause; but the Darwinian view (and that of Eldredge and Gould although they do not like to say so), that the gaps of the fossil record are due to its incompleteness, is scarcely discussed. The gradualism of Darwin, which applies not to rates of evolution but to the evolu-

tion of complex adaptation, is misunderstood by all the authors.

They have other objections too. For instance, they have all learned, from the popularizations of Sir Karl Popper and others, that natural selection is a circular argument (although Denton has still to learn from Pitman that Popper has "to some degree relaxed his position in a way that favours Darwinism"). And Denton thinks that natural selection cannot explain homology, including serial homology which "cannot by any stretch of the imagination be explained by descent from a common ancestor".

In the case of the two main objections I have dealt with, the authors have neglected the answers put forward by Darwin and his followers. Why should this be? Maybe they are unaware of them. Maybe they think only the devil needs an advocate. The texts, however, point to a third reason. The typical Darwinian of these pages is a man incapable of rational argument. Neo-Darwinism (we are told) is an "article of faith" (Pitman), a "belief-web" (Reid), a "paradigm" of "illusion" and "dogma" (Denton). It is believed in, and defended against critics, only because of "emotional commitment" and "extra-scientific factors", which are to be understood by the "sociology of knowledge". When the dogmas of neo-Darwinism are threatened, for instance by neo-Lamarckism, its fanatical adherents do not react with rational scepticism, about dubious factual claims. They remorselessly persecute the poor heretics, until they are driven into suicide or emigration, and

then trumpet their triumph in a ritual of collective orthodoxy. The doctrine of neo-Darwinism is thus seen as establishment cant, not rational science, and therefore hardly merits serious consideration. The critic may save himself the trouble, and ignore it.

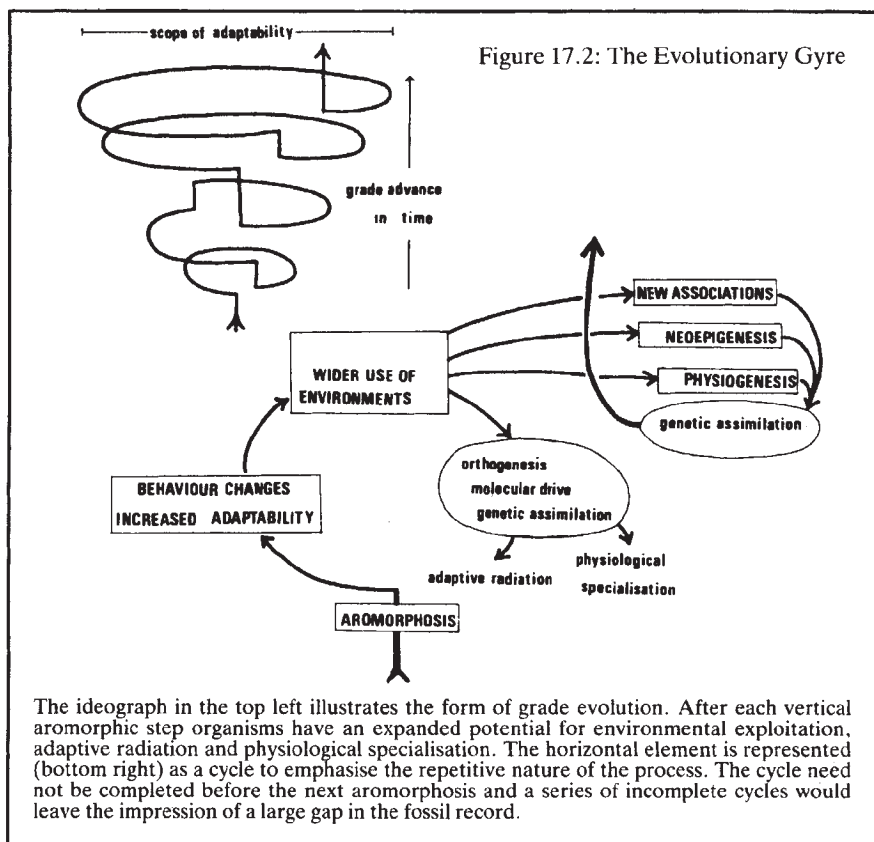
Reid is a rather different kettle of fish. His purpose is not, for the most part, destructive. He does include a chapter on some of the "difficulties" in Darwinism, and he notices others from time to time. The main problem, he believes, is that Darwinism is "reductionist" and has therefore "shied away from explaining the emergence of meaningful wholes from individually meaningless parts". It also suffers from two previously unnoticed defects: a new kind of circularity, and the "homeostasis paradox" ("if homeostasis is characterized as constancy, and evolution characterized as change, how did the homeostatic condition evolve?"). The circular argument is this:

Descent by modification through natural selection was Darwin's slogan. This linkage of two independent ideas was the core of Darwin's thesis and claim to originality, although it involved the circular argument that evolution occurs because of natural selection, therefore natural selection must operate; natural selection operates, therefore evolution must result.

Reid does not name names here, which is a pity because I cannot think of anyone who has fallen into this error.

But Reid's main purpose, as I have said, is positive. He was woken from his Darwinian slumbers by the student rebellions of the late 1960s and a reading of Thomas Kuhn. He duly discovered an alternative paradigm, of holists, and it is his aim in his book to introduce us to them. He has arranged people such as Driesch, Woodger, Bergson and von Bertalanffy into convenient groups; for each he gives some biographical details and summarizes their attitudes. They are in favour of emergence, integrative hierarchies and all that sort of thing. Anyone who is looking for an introduction to these people may find the book useful. But no one else will, because it is neither historical nor argumentative; it does not put their opinions in context nor does it try to show us that they were right.

Reid ends by expounding his own holistic alternative. It defies short summary, but fortunately he illustrates it with some figures, the most important of which is reproduced (left). Reid does not like neo-Darwinians, who do not treat all his friends and their enthusiasms with sufficient gravity. But now we have an opportunity to set things right. We should add the "evolutionary gyre" to our introductory lectures, expose our students to the alternative paradigm, and treat it with the seriousness it deserves. □



Mark Ridley is in the Animal Behaviour Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Holding back on dangerous play

Lawrence Freedman

Superpower Games: Applying Game Theory to Superpower Conflict. By Steven J. Brams. Yale University Press: 1985. Pp. 176. Hbk \$31.50, £22.50; pbk \$9.75, £6.95.

THERE was a time when game theory was all the rage in the strategic studies community. This was the "golden age" of strategic studies — the late 1950s and early 1960s — when it was fresh with new ideas and insights on how best to deal with the nuclear dilemmas that were perplexing policy-makers.

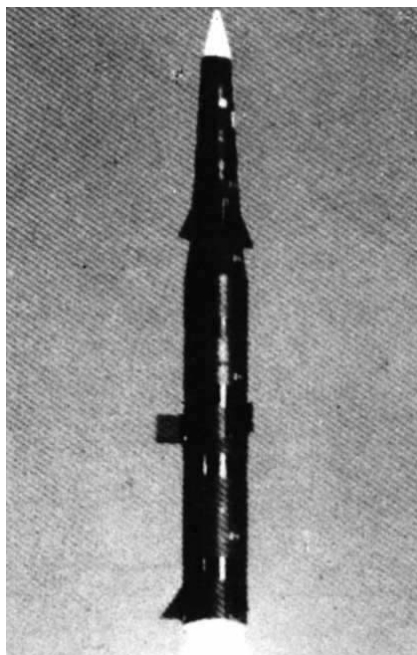
Game theory was ideal in two ways. First, it offered a methodology of some sophistication with which to challenge the traditional approaches to strategic issues based on history and politics and military experience. All of these were now deemed to be irrelevant in the nuclear age. Second, the assumptions behind game theory fitted in with the predispositions of the new generation of strategists. It was not that they saw superpower conflict as a "game"; the term only really referred to the idea of structured situations in which the ability of either "player" to achieve his best outcome depended on decisions taken by the other.

Superpower conflict at the time seemed to be suited to game theory for three reasons. First it involved two sides who were deeply antagonistic and likely to be so for the foreseeable future: there was no risk of the game dissolving in an outbreak of peace and harmony. Second despite this, the antagonism was incomplete. There was some common ground — to take the most obvious, a mutual interest in the avoidance of an outbreak of nuclear war. Third, as on the one hand the two sides were engaged in a contest for power and influence in the world, while on the other a nuclear war in which both would be devastated was the most likely result of any attempt to resolve differences through military means, it was useful to have a tool for identifying the solutions that logic might lead both powers to adopt in particular types of situation. If some solutions pointed to disaster then corrective action would need to be applied.

Within its limits game theory was quite valuable. It served as a means of emphasizing the extent to which superpower confrontation did not necessarily mean that one gained what the other lost (zero sum-games) but that there could be circumstances in which both won or both lost (non-zero sum games). It drew attention to the extent to which in non-zero-sum games both sides might tend towards the best of the worst solutions (minimax). Following through the logic of certain games

could illuminate aspects of the real dilemmas that the superpowers faced. But the approach did have its limits: the individual games had to be simplified in order to keep them manageable and a presumption of rationality had to be made. Moreover, the results of the games depended critically on the assumptions made as to the choices available to the two sides, and the values attached to the alternative outcomes (payoffs).

Accordingly the best exponents of game theory (notably Tom Schelling) never allowed themselves to be restricted by the theory and were often at their most effective in using it only as a starting point for a broader analysis. Those who had dabbled with it when fashionable soon lost in-



Pershing II missile — large chip in a deadly game.

terest. Nevertheless a number of scholars have persisted with the development of the technique.

Steven Brams is one such scholar who has worked hard to extend the range and applications of the basic theory. This book, which is made up of previously published articles, updated and reworked to give them some coherence, contains the results of his work. He is an unqualified believer in its continuing relevance. In the concluding chapter he states: "The evidence seems to me incontestable that game theory captures a substantial part of the logical basis and empirical reality of superpower conflict and provides a unified framework for its analysis" (p. 145).

This case can be assessed by examining the quality of the analysis and then the value of the conclusions. For the first, it has to be said that Brams makes few concessions to his reader. He does not attempt to justify his methodology against the standard objections or to introduce it for a lay reader. The mathematics is hard work for those (such as myself) who are

out of practice, and the argument is tight. So all that can be said, with the utmost humility, is that it seemed fine to me.

Much of the development of theory in this book constitutes an attempt to catch up with the empirical reality, and this accounts for much of the complexity of the argument. One means by which this is done is to test the theory against case studies such as the Cuban missile crisis of 1962 and the nuclear alert during the concluding stages of the 1973 Yom Kippur War. One concept of which Brams appears to be especially proud is that of "nonmyopic calculation" which essentially shows (if I have understood it correctly) how a different outcome is achieved when the belligerents are encouraged to take a long-term view and are not dominated by short-term considerations.

The question, however, is whether empirical reality is driving the theory to explain in complicated ways what might be explained more simply by more straightforward methods, or whether the theory can tell us something new. At times it is indeed illuminating but not to anything like the extent the author appears to believe.

In the end Brams suggests that the opponents act the way they do because of the "rules of the game" rather than because of anything very fundamental in the two sides' make-up. This seems to me to be going too far. It may be that the logic of the situations in which "players" find themselves influences the character of their choices, and the way those choices are made cannot be understood solely by reference to the rules of the game. Moreover, in structuring the games themselves, Brams makes some rather arbitrary judgements on the nature of the choices. For example, over Cuba, the options given for the United States are to blockade or to strike at the Cuban bases, and for the Soviet Union to retain the missiles or withdraw them. This does no justice to the actual range of choices — the United States might have done nothing; the Soviet Union might have escalated matters to take in Berlin. In the three basic issues Brams tackles — deterrence, arm-races and verification — only one dimension of the problem is addressed and the way that the problem itself is defined is controversial. On the whole the proposals at the end of the book are sensible, which gives one greater confidence in the analysis, but they are by no means wholly original and could have been arrived at by simpler methods.

In short Brams has demonstrated his skill as an exponent of games theory, and its continuing if modest value as an aid to working through the logic of superpower conflict. He has been far less successful in justifying the somewhat grandiose claims that he makes for the enterprise. □

Lawrence Freedman is Professor in the Department of War Studies, Kings College London (KQC), Strand, London WC2R 2LS, UK.

Checks on the spread

Gordon Thompson

Nuclear Proliferation Today. By Leonard S. Spector. *Vintage, New York: 1985. Pp.478. Pbk \$5.95.*

Non-proliferation: The Why and the Wherefore. Edited by Jozef Goldblat. *Stockholm International Peace Research Institute/Taylor & Francis: 1985. Pp.343. £11, \$19.*

Safeguarding the Atom: A Critical Appraisal. By David Fischer and Paul Szasz. Edited by Jozef Goldblat. *Stockholm International Peace Research Institute/Taylor & Francis: 1985. Pp.243. £16, \$28.*

Most people would agree that the world will be safer if membership of the nuclear weapons club does not grow. Likewise, most seem to think that we would all be better off if the present members were to stop their frenzied arms race.

Thus we are faced with a problem of nuclear proliferation in two basic dimensions — "horizontal", which refers to the acquisition of nuclear weapons by additional countries, and "vertical", which refers to increases in the sizes or capabilities of existing nuclear arsenals. (We owe this terminology to Indian diplomats who, in the 1960s, coined the term "nuclear proliferation" specifically to link the two dimensions.) The three books reviewed here, if taken together, provide an up-to-date and comprehensive picture of the state of nuclear proliferation, although only in its strictly horizontal dimension.

With Leonard Spector's book, the Carnegie Endowment for International Peace (based in Washington, DC) has launched a series of annual reports on the spread of nuclear weapons. *Nuclear Proliferation Today* provides an account of the proliferative activities of eight "problem" countries — Argentina, Brazil, India, Iraq, Israel, Libya, Pakistan and South Africa — together with discussions of the role of China, and of the status of controls and safeguards on nuclear technology. The eight nations, using a variety of approaches to getting the necessary materials and technology, either have nuclear weapons already or are seeking to acquire them. India publicly acknowledged its nuclear test in 1974, although labelling it a peaceful explosion. Israel, which may or may not have conducted a test, is generally believed to have a few tens of nuclear weapons, perhaps kept in a disassembled form in times of peace. The other six nations are in more ambiguous positions or are less advanced in their quest for a Bomb.

Two of the eight countries (Iraq and Libya) have signed the Non-Proliferation Treaty (NPT), although they apparently do not feel bound by its provisions. The others have refused to sign, generally

arguing that the NPT is discriminatory and seeks to perpetuate the dominance of the nuclear powers.

Nations have had a variety of motives for their decisions to sign or to stand aloof from the NPT. For representative groups of countries, these motives are explored in *Non-proliferation: The Why and the Wherefore*, a collection of essays describing the domestic, international and technical factors which brought each country to its decision. The group of NPT non-signatories studied includes two nuclear-weapon states (China and France), all the non-NPT "problem" countries mentioned above, and Spain. Representing the NPT signatories is a group of six countries ranging from Canada (which has never sought an independent nuclear arsenal) to Taiwan (where plans to build a reprocessing plant were abandoned following heavy pressure from the United States in 1976).

For the book, the Stockholm International Peace Research Institute (SIPRI) has assembled a stable of well-qualified authors (many of them citizens of the countries they discuss) and, to assist in framing the conclusions presented in the overview chapter, the Institute held a workshop in October 1984, attended by a separate group of international experts. Also, SIPRI has drawn on its own resources to prepare several detailed appendices, including one which provides a thorough account of the nuclear infrastructure in each country considered (although the data for China and France are badly out of date).

Comparatively speaking, the conclusions of both SIPRI's study and of Spector's more narrowly focused book are mildly encouraging; that is, the present pace of horizontal proliferation is slow compared with the pace of vertical proliferation. In part, this is because of the emergence of a broad international consen-

sus that the nuclear club should not grow. From this has flowed both moral pressure and the ability to impose an international non-proliferation regime which has complicated the task of any nation seeking nuclear capability. Moreover, the financial costs and technical problems of acquiring an arsenal of nuclear weapons and appropriate delivery systems have proved formidable for developing nations. Finally, growth of the nuclear power industry has been much slower than anticipated; in consequence, possession of a complete nuclear fuel cycle (with the concomitant possibility of diverting fissile materials from reprocessing or uranium enrichment plants) cannot be economically justified for any country, particularly a developing nation.

Central to the non-proliferation regime is the safeguards system operated by the International Atomic Energy Agency (IAEA). This system is the main subject of another SIPRI book, *Safeguarding the Atom: A Critical Appraisal*, which has as its principal author the former IAEA official David Fischer. It also contains a section by Paul Szasz (also formerly with IAEA) and other sections and detailed appendices prepared by SIPRI.

The safeguards laid down by the IAEA are supplemented by other arrangements. First, the Treaty of Tlatelolco, which seeks to make Latin America a nuclear-free zone, is based upon IAEA regulations but also provides for special inspections by a regional agency in the event of suspicious activities. Secondly, the European Economic Community has its own internal system of EURATOM safeguards, which rather wastefully duplicates that of the IAEA. Then there is the Nuclear Suppliers' Group, often known as the London Suppliers' Club, a group of those nations which dominate the export of nuclear materials and technology and which have agreed not to supply certain items

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

MASSON

Introduction à une biologie des populations

par J.-M. LEGAY et D. DEBOUZIE

1985, 152 pages
2-225-80619-5, 119 FF

The only book, written in French, that treats of population matters, in terms of structure, sociology, genetics and biology.

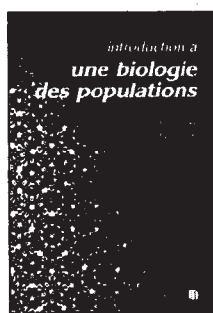


Table of contents

Les structures de socialité. Les structures spatiales. Les structures temporelles. Les structures génétiques. Les structures de parenté. Conclusions générales. Index.

Membranes biologiques

Structure, transports, bioénergétique
par E. SHECHTER

1984, 232 pages
2-225-80367-6, 145 FF

This textbook describes three main aspects of biological membranes - structure, transport, bioenergetics - which control most of their other functional properties.

Table of contents

STRUCTURE ET DYNAMIQUE STRUCTURALE. Lipides membranaires. Conformation des chaînes hydrocarbonées des lipides. Fluidité membranaire. Diffusion latérale des lipides et des protéines. Mobilité et asymétrie transversales des constituants membranaires. Protéines membranaires: extraction, structure, reconstitution.

TRANSPORTS. Rappels de thermodynamique en relation avec le transport. Diffusion passive non ionique. Transport de l'eau. Diffusion ionique. Différence de potentiel électrique transmembranaire. Diffusion facilitée. Transports actifs.

BIOÉNERGÉTIQUE MEMBRANAIRE. Le concept chimiosmotique. Rappels d'oxydo-réduction. Chaîne de transfert d'électrons (membrane interne mitochondriale). Premier couplage entre transfert d'électrons et transport de protons (membrane interne mitochondriale). La différence de potentiel électrochimique de protons. Deuxième couplage entre diffusion facilitée de protons et synthèse d'ATP (membrane interne mitochondriale). Autres systèmes membranaires siège des conversions d'énergie.

For further information, write to
MASSON Publisher,
120, boulevard Saint-Germain,
75006 Paris, FRANCE.

unless the recipient nations agree to safeguards (the list of sensitive items is similar to the so-called "Zangger Trigger List", drawn up by a committee of nuclear suppliers chaired by Claude Zangger of Switzerland). Finally, there are the arrangements made by individual countries, of which the most important is the United States Nuclear Non-Proliferation Act of 1978.

The NPT and IAEA safeguards, together with these other arrangements, constitute the non-proliferation regime. As mentioned above, the regime has been helpful in restraining horizontal proliferation: much of the nuclear industry outside the nuclear weapon states is under safeguards, as described in detail by David Fischer, while the "problem" countries generally find it difficult to obtain materials and technology for their weapons programmes. Yet, the future stability of the regime is in question.

As both SIPRI books acknowledge, without pursuing the point, the non-proliferation achievements of the past 20 years could be swept away by changing circumstances. First, the NPT remains in force only until 1995 and it is by no means clear that the treaty will be renewed or replaced by a better arrangement at that time. Many signatories are dissatisfied with the failure of the nuclear-weapon parties (the United States, the Soviet Union and Britain) to live up to their side of the bargain — that is, to end their nuclear arms race. Secondly, regional instability could precipitate a rush of nuclear competition between certain powers. Pakistan and India, for example, could fall prey to such competition, and the tension between these two nations illustrates some important linkages of proliferation. American and Soviet nuclear competition has fed the Chinese nuclear build-up. In turn, this has contributed heavily to India's flirtation with the Bomb. Pakistan has reacted, and thereby fed the fears of Israel about an Islamic Bomb. To compound these linkages, China has apparently assisted Pakistan in its quest for nuclear weapons.

An unavoidable conclusion of all this is that nuclear proliferation cannot be considered in just one dimension. Equally, it is hard to envisage any substantial improvement in the horizontal dimension without progress on control of vertical proliferation. In this respect, all three books are deficient. In particular Leonard Spector reflects the view that proliferation is something which other nations do. This attitude, common in the United States, is perhaps understandable given America's role in constantly raising the qualitative level of vertical proliferation. While SIPRI, as befits its fine record of analysing military activities of all kinds, provides a more balanced view, *Non-proliferation: The Why and the Wherefore* does not really pursue the matter of cross-dimensional linkages.

Readers seeking to understand vertical proliferation have the choice of a vast literature, but at some point would be well advised to consult William Arkin and Richard Fieldhouse's *Nuclear Battlefields: Global Links in the Arms Race* (Ballinger, 1985) which provides a succinct summary of how the nuclear confrontations of the five major powers proceed in a myriad of ways. By outlining the vast global infrastructure which supports development, testing and deployment of nuclear weapons, the book makes clear that the implications of vertical proliferation reach further than we might think.

Two examples are instructive. First, through their participation in the Eastern and Western blocs, NPT signatories can function almost as though they had never signed the treaty. Indeed, six NATO countries which are supposedly "non-nuclear" by virtue of being NPT signatories, and four Warsaw Pact countries in the same position, have numerous delivery systems certified to fire nuclear warheads. When needed, the weapons will be released to them by their nuclear-armed allies. Secondly, the intelligence, targeting, basing, training and communications infrastructure for the nuclear arsenals is spread around the globe, taking in many countries which fancy themselves "nuclear-free" and making a mockery of the idea of nuclear-free zones.

If the non-proliferation edifice begins to crumble, we will rapidly become aware of a third dimension of proliferation. This, the "latent" dimension, refers to the growing ability of many nations rapidly to field a nuclear arsenal should they so decide. The necessary materials and expertise are becoming more generally available as a result of nuclear power programmes. In particular, the wide adoption of nuclear fuel reprocessing and the development of new enrichment technologies imply the stockpiling of large quantities of fissile materials. Despite the lack of economic justification, industrial nations, including Japan and West Germany, are pressing ahead in these areas. Since appropriate nuclear weapon delivery systems are readily available, such nations could field large arsenals with surprising speed. But again, this possibility is discussed in none of the three books reviewed here.

Viewed from a wider perspective, then, nuclear proliferation is a far more serious problem than the strictly horizontal trends of recent years might suggest. It is not alarmist to say that we urgently need to tighten international controls on all forms of proliferation. The current non-proliferation regime provides a basis for doing this, but the world community needs to find the vision to adopt substantial reforms. □

Gordon Thompson is Director of the Institute for Resource and Security Studies, 27 Ellsworth Avenue, Cambridge, Massachusetts 02139, USA, and coordinator of the IRSS Proliferation Reform Project.

Towards a New World synthesis

Colin A. Russell

Chemistry in America, 1876–1976. By Arnold Thackray, Jeffrey L. Sturchio, P. Thomas Carroll and Robert Bud. Reidel: 1985. Pp.564. Dfl.210, \$79.50, £53.50.

A CASUAL reader who decides to browse through this volume is in for quite a surprise. From the title one might expect an account of American chemistry (here meaning chemistry in the United States) on the lines of those engaging and once-popular series "a hundred years of this or that": bland, popular, anecdotal, synoptic, well-illustrated, easy on the eye and reasonably analytical. For the price and size one might expect all this on a lavish scale. Should our browser dally even for a moment at the title page he will, however, receive the first hint of unpredictability. The subtitle "historical indicators" suggests uncharted seas ahead. For, if these are not to do with the litmus-type indicators familiar to all chemists, what on earth are they? The confusion is not resolved by the apparently perverse and deliberate appropriation of the very phrase "chemical indicators" that has for so long had a quite precise if pedestrian connotation of an entirely different kind.

"Indicators", as used by these authors, are long-running trends in a series of statistical data. They are measurements of change. If, for example, you analyse census and other information about the employment of chemists in the United States, you find for the past century or so an annual growth rate of about five per cent in the chemical work-force, and that chemists have increased from 2 to 15 per 10,000 of the total numbers employed in any work. On the other hand, the proportion of high school chemical students who later specialize in chemistry has declined steadily, as has the esteem accorded to chemistry in public opinion surveys and its coverage in newspapers. And since about 1910 there has been absolute growth *but relative decline* in the place of chemistry in the internal economy of American universities.

In six chapters, occupying less than half the book, the authors present a profusion of tables, graphs, bar-charts and the like to show how American chemistry has changed in the context of industry, education and its professional organization. Their conclusions are helpfully summarized in each chapter by a few "indicator highlights". For people who are prepared to take the arguments on trust these brief conclusions are probably all that matter: two or three pages of print altogether. However with typical trans-Atlantic thoroughness the authors will not allow

us off the hook so easily. They parade their hard-won information in a spirit of modesty and self-criticism, and delight in warning of the hazards and limitations of their approach. As if to hammer the point home they append to the book over 200 pages of tables, some appendices on methodology and a bibliography of massive proportions.

The methods, like the idea of "indicators", derive from the social sciences. So also, apparently, does the phraseology, though this is mainly evident when the authors are being self-consciously coy about their strategy. Indeed it seems that the most avid users of the book are likely to be those students of society for whom chemistry is but another manifestation of social conditioning and manipulation. Such enquirers will find a wealth of raw material brought together for the first time, and also many hints and ideas for future research. Yet it is as part of a series devoted to the history of chemistry that this book appears, so it is not unreasonable to enquire what chemists and historians of science might make of it.

Chemical readers will here find little of the *content* of their science, even historically treated. They will, however, discover a great deal about its *context*, in which only the most myopic benchmarker will be altogether uninterested. The conclusions, at least for American chemists, are not wholly agreeable, for the subject's decline is remorselessly depicted in the many graphs and tables. For those who plan and teach, these trends, however depressing, cannot now be ignored. If, as the authors claim, "chemical education begins to look like a sieve, leaking credentialled chemists in all directions", one would have thought that those who advise American youth would at least be glad to know. So, one imagines, would those responsible for science policy.

For historians of chemistry, the book will have little appeal to those who like their history warm, human, personalized and cosy. Nor will it be of much use to their polar opposites, historians of scientific ideas, philosophies and techniques. People to whom quantitative historical science is a contradiction in terms ("history by numbers" such efforts are contemptuously called) will look elsewhere for inspiration. Nor must the sceptics be dismissed as mere backwoodsmen. They have a real point when they urge that huge compilations of statistics either prove what you want them to prove (just select the right parameters) or that they merely clothe the blindingly obvious with pseudo-quantitative respectability. It is to the great credit of the authors of this book that they are well aware of these dangers, and do not attempt to conceal them and can (in a couple of apposite cartoons) even laugh at themselves.

Those whose interests lie in the so-called social history of science will find in *Chemistry in America* a fund of good ideas

and useful information, especially if they are concerned with institutions or education. Yet the book also has drawbacks for such people. The material relates to only one country — though admittedly a world leader in chemical research — and one is bound to seek comparisons with Britain, Germany and other countries if only for the earlier part of the period under consideration. Moreover, it is important that not only should national trends be compared but that allowance be made for different meanings for the variables measured. Thus the criterion of professionalization in chemistry is a very American one and different from that in Britain where the phenomenon first arose. Similar caution must be given about the definition of various classes of chemists which mean different things in different countries; "assayer" is a case in point.

It is not difficult to criticize the selection and presentation of the data. One looks in vain for information on such topical issues as the changing role and numbers of women in chemistry, the relation between chemistry and war, and even the question of chemists' salaries. In education, trends in chemistry are compared with those in history rather than in physics and biology. Again, there are surely more important issues to record than the number of ex-chemists who become deans in American universities or the ages (high) of Presidents of the American Chemical Society. However, the prominence given to that specific institution is one of the particularly creditable features of a book that is always in danger of submersion in an ocean of generalities.

In reading the book I became gradually conscious of a sense of *déjà vu*. Somewhere I had met something very like it before, but where? The impression crystallized into recognition by about Chapter 3. Its predecessor was nothing less than the massive *History of Chemistry* by J.R. Partington, published by Macmillan in four volumes between 1961 and 1970. Superficially the two works are light-years apart yet fundamentally they are doing the same thing: presenting a highly idiosyncratic view of a mountain of information with little or no attempt to digest it into a coherent, readable overview of the subject. Here, perhaps, is a Partington of a new era, with all the weaknesses and strengths of the original. Indigestible it may be, and one could certainly not read it from cover to cover. But if it turns out to be half as useful as Partington in supplying raw information and in catalysing future work, it will earn the gratitude of generations to come. Whether the trends it heralds in chemical history — empirical, pragmatic and semi-quantitative — will prove to be fruitful, time alone will tell. □

Colin A. Russell is Professor of History of Science and Technology and Director of the History of Chemistry Research Group, Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Space shuttles

Peter Kemp

Contact: *A Novel*. By Carl Sagan. Simon & Schuster: 1985. Pp.432. \$18.95.

AMONG the numerous — and often numinous — quotations strewn around Carl Sagan's latest novel, is Einstein's claim that "the cosmic religious feeling is the strongest and noblest motive for scientific research". True to this belief, Sagan's saga of galactic exploration keeps juxtaposing the scriptural and the scientific: apocalyptic sects and astronomers, angels and astronauts, messiahs and supernovae, tales of mystic experience and talk of transcendental numbers. Even in shape — opening with genesis and ending with revelation — his book seems to have its eye on the Bible.

The first page records the birth of Ellie, a baby who comes into the world with a look of "puzzlement". As she grows up, this matures into intellectual curiosity — especially about possible extraterrestrial life. Finding her vocation as a radioastronomer, Ellie pioneers new techniques for scanning the heavens — and, as the second millennium approaches, her hunch that there are powerful beings out there proves justified. From the vicinity of Vega, a coded communique is picked up, transmitting instructions for the manufacture of a massive machine. Despite objections from some quarters that this is a snare of Antichrist, construction work goes ahead. At last — on 31 December, 1999 — Ellie and four international colleagues board what has turned out to be a spaceship ready for launching.

Up to this point, Sagan's story has kept to fairly well-trodden ground, such as the way space research might foster "transnational" thinking and global unity: a notion regularly mooted in science fiction since the days of H. G. Wells. The crucial test for space travel stories, though, is what happens when their characters zoom away from Earth. After take-off, the narrative is at maximum risk of nose-diving into the banal.

This is a hazard Sagan doesn't altogether steer his way around. Though he musters an imposing array of technical terminology — as his crew shuttle through space, they exchange arcane remarks about "time-like Killing vectors, non-Abelian gauge invariance, geodesic refocusing, eleven-dimensional Kaluza-Klein treatments of supergravity" and the like — he also employs analogies that seem something of a let-down. The mysterious black tunnels along which his starcraft is whisked are, it is improbably suggested, like "a subway"; star systems loom up like stops along the line; the gigantic junction at the heart of it all is "Grand Central Station".

Even more surprisingly down-to-earth

is the destination at which his travellers disembark: a kind of far-fetched Disneyland. Alighting on a simulation of a tropical beach — all palm trees and balmy breezes — they are greeted by aliens who have considerably assumed not merely human shape but that of their guests' "deepest loves". More suggestive of *When You Wish Upon A Star* than of what you might expect to find when you travel to one, these sentimental look-alikes are also designed, it appears, to figure as equivalents of scriptural divinities: "beings who live in the sky, beings enormously knowledgeable and powerful, beings concerned for our survival". "How . . . theological

. . . the circumstances had become" Ellie reflects excitedly. But, for all the book's straining towards some lofty spiritual sphere, its final message is mundane. Returned to Earth after her galactic gallivantings, Ellie clutches the insight that "For small creatures such as we the vastness is bearable only through love". A piece of enlightenment you'd hardly think it necessary to travel light-years to acquire, it's disappointingly typical of the way this story tends to drop from the nebular to the nebulous. □

Peter Kemp is an Associate Lecturer in English at the Middlesex Polytechnic, All Saints, White Hart Lane, London N17 8HR, UK.

A light on the past

Hendrik B.G. Casimir

History of Physics: Readings from Physics Today. Edited by Spencer R. Weart and Melba Phillips. American Institute of Physics: 1985. Pp.375. Pbk \$25.

WHEN, in 1948, the American Institute of Physics (AIP) launched *Physics Today*, it did not intend the journal to deal with the past. The new publication addressed itself to members of the AIP and of its several affiliated bodies, and aimed at keeping them informed about what was happening in physics, both from a scientific and from an institutional point of view. Only in 1952 did a contribution of a historical, or at least of a retrospective character appear in its columns, but from then on historical articles became a more or less regular feature. *History of Physics* — a somewhat grandiose title which is put into proportion by the inconspicuously printed subtitle, *Readings from Physics Today* — is an anthology of such articles, together with a brief introduction.

The introduction gives a useful survey of the activities of the AIP in the field of history and of the development of other such work in the United States since the Second World War. In another respect, however, it is not entirely satisfactory; when the editors state that there was hard-

ly any organized research into the subject before 1950 — that there were no books, no papers, no journals on the history of physics — they are taking a rather parochial view. On this side of the Atlantic we had E.T. Whittaker's great work *History of the Theories of Aether and Electricity* and the famous *Ostwald's Klassiker*, annotated reprints of important papers, to name but two examples. Moreover there were many biographies and autobiographies, and several journals dealing partly or entirely with the history of science, including physics.

This, however, is a point of minor importance and should not deter prospective readers. What really matters is the selection of material and in this respect the editors have done an excellent job. Not surprisingly, the articles differ widely in style and in scope, which makes it difficult to characterize the book as a whole. The best I can do is to discuss a few typical examples.

Two papers, one by Goudsmit and one by Uhlenbeck, deal with the birth of the notion of the spinning electron and with the circumstances that surrounded it. Here we have the story of a seminal idea, related by the originators half a century later. In consequence not only do these essays give a vivid impression of quantum physics and spectroscopy in the 1920s, and of the little group working with Ehrenfest at Leiden University, but they also provide an intriguing insight into the later

264

NATURE

[FEBRUARY 20, 1926]

Letters to the Editor.

[The Editor does not hold himself responsible for opinions expressed by his correspondents. Neither can he undertake to return, nor to correspond with the writers of, rejected manuscripts intended for this or any other part of NATURE. No notice is taken of anonymous communications.]

Spinning Electrons and the Structure of Spectra.

So far as we know, the idea of a quantised spinning of the electron was put forward for the first time by A. K. Compton (*Journ. Frankl. Inst.*, Aug. 1921, p. 143), who pointed out the possible bearing of this idea on the origin of the natural unit of magnetism. Without being aware of Compton's suggestion, we have directed attention in a recent note (*Naturwissenschaften*, Nov. 20, 1925) to the possibility of applying the spinning electron to interpret a number of features of the quantum theory of the Zeeman effect, which were brought to light by the work of the German physicists, Sommerfeld, Paschen, and

this moment of momentum is given by $K\hbar/2\pi$, where $K = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$. The total angular momentum of the atom is $J\hbar/2\pi$, where $J = 1, 2, 3$. The symbols K and J correspond to those used by Landé in his classification of the Zeeman effects of the optical multiplets. The letters S, P, D also relate to the analogy with the structure of optical spectra which we consider below. The dotted lines represent the position of the energy levels to be expected in the absence of the spin of the electron. As the arrows indicate, this spin now splits each level into two, with the exception of the level $K = \frac{1}{2}$, which is only displaced.

In order to account for the experimental facts, the resulting levels must fall in just the same places as the levels given by the older theory. Nevertheless, the two schemes differ fundamentally. In particular, the new theory explains at once the occurrence of certain components in the fine structure of the hydrogen spectrum and of the helium spark spectrum

k old K new J

Birth of the notion of the spinning electron

views of two senior scientists.

A delightful curiosity is the text of an address given by Einstein at Kyoto University on 14 December 1922, "How I Created the Theory of Relativity". The lecture was given in German and a Japanese physicist, J. Ishiwara, provided a translation. In 1923 Ishiwara published his Japanese text and in 1982 *Physics Today* printed an English translation by Yoshimasa A. Ono. Clearly, knowing that the lecture was to be translated, Einstein had formulated his ideas as concisely and simply as possible, but even so it is remarkable to see how well the text has survived two consecutive translations. There are several other articles in which the story of an important discovery or invention is told by people who were not necessarily the only originators but who did play an important part — Frisch and Wheeler on the discovery of fission is a good example, as is the history of the cyclotron told by Livingston and McMillan.

Also included are several biographical essays. Most of these are written by authors who were well-acquainted with their subjects, as their students or collaborators, which makes the stories lively and readable — here, Oliphant's "The Two Ernests" (Rutherford and Lawrence) is especially illuminating. Quite different again are the scholarly papers on specific aspects of the history of physics, for example Martin Klein's thorough study "Thermodynamics and Quanta in Planck's Work". Klein shows how Planck, a conservative rather than a revolutionary, found it hard to reconcile himself to his own innovation, but finally arrived at the conclusion "that the quantum of action played a far more significant part in physics than I had originally been inclined to suspect".

It is understandable that the contributions on institutional history deal almost exclusively with American establishments, and the same is true, though to a lesser extent, of the articles grouped under the heading "Social Context". But this does not mean that they are of no interest to non-Americans, and in many cases the problems discussed are international in nature. Of particular interest here is the article by Vera Kistiakowski with the provocative title "Women in Physics: Unnecessary, Injurious and Out of Place?".

Few people will read this book from cover to cover — the editors even warn against trying to do this. But all physicists and many non-physicists will enjoy leafing through it, will admire the many excellent illustrations, and will sooner or later start reading in earnest. □

Hendrik B.G. Casimir, *De Zegge 7 5591 TT Heeze, The Netherlands, studied theoretical physics at Leiden, Copenhagen and Zürich during the 1920s and 1930s and later became joint director of research at the Philips Company, Eindhoven. His autobiography, Haphazard Reality: Half a Century of Science, was published by Harper & Row in 1983.*

In the realm of the cosmocrats

Edward Harrison

The Cosmological Distance Ladder: Distance and Time in the Universe. By Michael Rowan-Robinson. W.H. Freeman: 1985. Pp.353. \$35.95, £19.95.

COMTE de Buffon once said that the forms of life seem more like chain mail than a Great Chain of Being. Similarly the measure of the universe seems more like a network of ladders than a single cosmological distance ladder. (An irreverent person might say more like a game of snakes and ladders.)

The scale of the universe has been increased several times since Hubble wrote *The Realm of the Nebulae* in 1936. The Hubble term H in the velocity-distance law (velocity = $H \times$ distance) has decreased in value by an order of magnitude and now stands at either 100 or 50 km s⁻¹ Mpc⁻¹. The de Vaucouleurs school champions the higher value and the Sandage-Tammann school the lower value. The reciprocal of H is about fifteen thousand million years and serves as a rough measure of the age of the universe.

Rowan-Robinson unifies his splendid and comprehensive survey of the universe — planets, stars, star clusters, galaxies, galaxy clusters, superclusters — with the theme of the cosmological distance ladder. In the preface he writes that, at a conference in 1976,

Gerard de Vaucouleurs on the one hand, and Allan Sandage and Gustav Tammann on the other, arrived at estimates of the size of the universe, as measured by the Hubble constant, differing from each other by a factor of two. Moreover, when I asked the protagonists what was the range outside which they could not imagine the Hubble constant lying, these ranges did not even overlap. . . . It was to understand this disagreement that I set out to write this book.

Using numerous figures, tables and appendices that display a wealth of data, the author builds an array of distance indicators and presents a difficult subject in a methodical manner. Skilfully he adjusts, deftly weighs, and adroitly averages divergent results.

The first rung on the ladder is the distance between the Sun and the Earth. Then comes the distance to the Hyades, the nearest open star cluster, determined by parallax and moving cluster methods. The size of this preliminary and important step, now estimated to be 150 light-years, has grown 30 per cent since 1940.

We measure in Hyades-units the distance to other open clusters, applying various methods including evolutionary curve fitting. Then, with allowance for absorption of light by interstellar dust, we use Cepheids and RR Lyrae variable stars as yardsticks to gauge the distances of globular clusters in our Galaxy and of nearby galaxies in the Local Group (rival schools going different ways and making different allowances for extinction of light). Novae, supernovae and HII regions, adjusted for extinction in the observed galaxy, serve as intergalactic

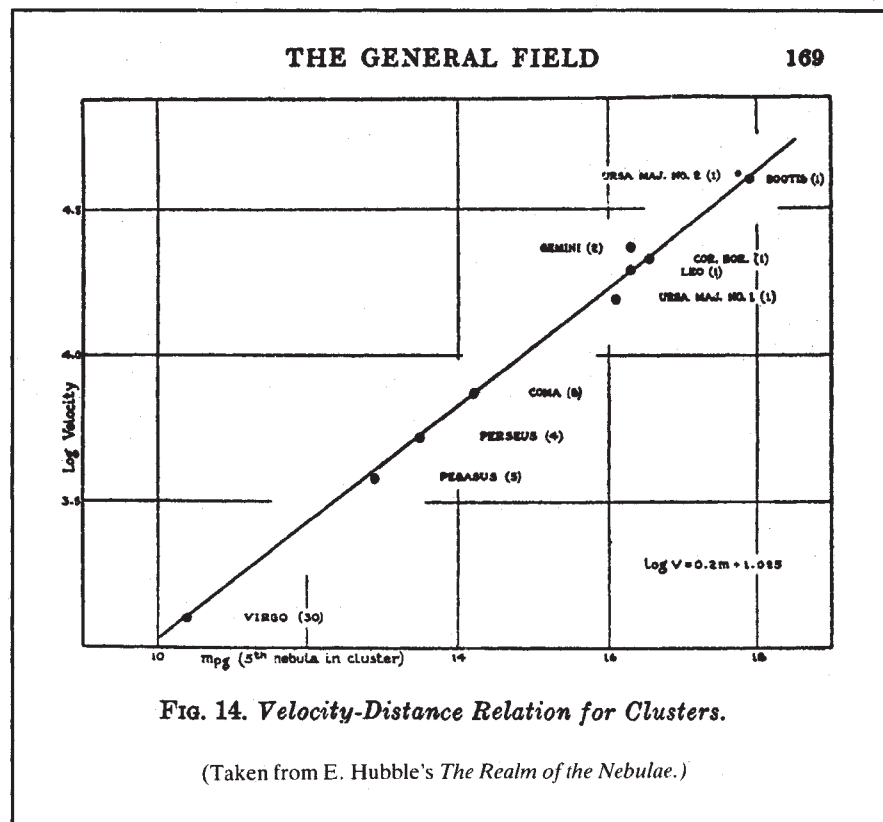


FIG. 14. Velocity-Distance Relation for Clusters.

(Taken from E. Hubble's *The Realm of the Nebulae*.)

indicators of distance; here the methods become less reliable and the uncertainties increase yet further. Supergiant galaxies then act as stepping stones to remoter regions.

It began with the Greeks. Rowan Robinson gives them a page or two and then jumps, as one expects in an astronomer's history of astronomy, from Aristotle to Copernicus. The Epicurean system (an infinite atomic universe) and Stoic system (an island cosmos surrounded by an infinite extracosmic void) receive scant attention, although both dominated the intellectual climate of the ancient world. The great historical drama of the mediaeval revival of the Aristotelian system, the eclipse of this system in the late Middle Ages by a renescent Stoic system, the rebirth of the Epicurean system, and the eventual overthrow of the Stoic system in the Great Debate of the 1920s also receives little recognition. We are told that in the seventeenth century Cassini measured the Sun–Earth distance to within 10 per cent of the modern value, but not that Newton, by using James Gregory's photometric method of comparing the outer planets with the brightest stars, determined with astounding accuracy (within a factor of two for Sirius) the distance of the nearest stars.

In a chapter entitled "The Cosmological Models", which is devoted to theoretical topics, the author discusses dynamic models, redshifts, proper and luminosity distances, the early universe, and the motion of the Solar System relative to the microwave background and to the Local Supercluster. Perhaps more such analysis in the other chapters would have helped the reader to surmount the mountains of data to be found in them.

The axioms or irreducible assumptions of modern cosmology rarely receive adequate attention, and the book would have benefited from some discussion of these hidden assumptions. The Hubble (and deceleration) term tell us how the universe expands. But how exactly do we need to know its value? Theorists first postulated the velocity–distance law, which applies at any instant throughout a homogeneous expanding universe. But when we look out in space we also look back in time, and homogeneity (unlike isotropy) cannot be directly seen. Observations made on the backward light cone are adjusted for evolutionary effects and mapped in a hypothetical homogeneous space. At best we verify that redshifts and estimated distances are consistent within the theoretical frame of a geometric and dynamic model that is itself unverified. Perhaps our models are too simple and we do not know enough cosmology to make an exact value of the Hubble term meaningful? □

Edward Harrison is Professor of Astronomy in the Department of Physics and Astronomy, University of Massachusetts, Amherst, Massachusetts 01003, USA.

Halley's comet in print

David W Hughes

ON 16 October 1982 Halley's comet was picked up using a charge-coupled device at the prime focus of the 5.1 m Hale telescope on Mount Palomar. It has been monitored ever since. As it approaches the Sun it brightens and will become just visible to the naked eye between December 1985 and April 1986. But there are two problems for observers. On 9 February 1986 the comet is at its closest to the Sun and for a six-week period around then is in too bright a region of sky to be easily visible. Also, between December and April it sweeps from the northern to the southern sky, requiring the northern observer to reciprocate by moving to the southern hemisphere. Unfortunately planet Earth does not get very close to the comet this time, so we are not going to be overawed by startling searchlight-like beams radiating up from the twilight horizon. In fact the mere act of catching sight of the comet will require some effort.

Halley's return to the inner Solar System has been greeted by a whirlwind of books, and thus a rather rare state of affairs as far as the reviewer is concerned. Most of them in fact say much the same things, use the same illustrations and are aimed at a like audience. Most are written by non-specialists who, for simplicity, have consulted only secondary sources of information and have consequently made many mistakes. Reviewing them is akin to essay-marking. To help the prospective buyer the books are divided into categories, with a star rating for each according to whether it is thought to be excellent, good, indifferent or poor. The five-star rating is reserved for those rare books in the pile that will "last", to be consulted around AD 2061 when Halley's comet returns once more.

The categories are as follows: (i) Books specifically designed to help the reader see the comet (the message being that you have a once-in-a-lifetime chance so make the most of it); (ii) guides for teachers, that is books specifically designed for the organizers of classes; (iii) specialist books; and (iv) general books.

Halley's Comet by Francis Reddy leads the first category. It abounds with good advice: "If you want to see the comet you had better prepare for it, familiarize yourself with the constellations through which Halley moves, get away from cities, use your binoculars and travel south". Twenty maps are provided, for three Earth latitudes (55°N, 40°N and 35°S). These not only show the comet and its tail against the starry background but also include (where appropriate) the Moon, Mars and Jupiter. A superb three-dimensional model of the orbits of the comet and Earth is an additional bonus.

Such models are a great help when it

comes to visualizing the movement of the comet with respect to the Sun and Earth, and Perihelion Scientific have produced one, *Halley's Comet Survival Kit*, which is 85 cm across. Another accompanying model enables you to calculate the altitude and azimuth of the comet throughout the November 1985–May 1986 period. Both are great fun to assemble and use.

If you are happier with a planisphere, David Chandler has designed a set of three (for northern latitudes 26°, 35° and 44°). These show the comet's path through the constellations and it is simplicity itself to calculate when and where it can be seen. Unfortunately there are no hints as to its brightness.

Halley's Comet Finder by Ben Mayer is aimed more at the astronomically initiated and gives a detailed review of the comet's journey across the sky. The section on astrophotography in this book is most useful.

For the complete beginner the editors of *Sky and Telescope* have produced *Mr. Halley's Comet* (I've no idea why they refer to a man who was an eminent professor and had two doctorates as "Mr."). This is an excellent and inexpensive guide marked only by the sparsity of stars on the sky map. The fact that a pair of binoculars has only a limited field of view is rightly stressed, but the difficulty of finding the comet and recognizing stars in the field is somewhat overlooked. The guide seems to be designed in the hope that the comet would be much brighter.

Moving into the classroom, both *Comet Halley Returns* and *Halley's Comet Activities Manual* get good marks and will provide science teachers and their scholars with hours of activity understanding comets, modelling their orbits, recording observations, computing examples of Keplerian motion, introducing the life and times of Edmond Halley and trying to appreciate his achievement. *Halley's Comet: Teacher's Notes* comes with posters and is biased towards young (9–14) or innumerate pupils (and supposedly innumerate teachers too!). It approaches cometary science from a sociological and refreshingly critical standpoint — I'm still considering "the possibility that women astronomers do not appear in astronomy books because those books are generally written by men", one of the topics suggested for classroom discussion. Mary Ashby, the author, has a clear grasp of what children find interesting and what they find difficult to understand, and has planned a series of exercises for explaining difficult concepts.

Nine books fall into the specialist category, two of which are bibliographies. Ruth S. Freitag lists 3,289 references to both popular and scientific cometary literature. The list is alphabetical by author and each

reference is followed by a short description of the item. This book is an indispensable source for anyone seeking data on the comet's previous apparitions, technical literature on the Halley spacecraft missions or manifestations of popular interest in comets. Just turn to the index, and look up, for example, "Essays and sketches, humorous"; "fiction"; "recovery (1759)"; "transits across the sun's disk"; "music"; and "mass spectroscopy". What a godsend this book is. I hope that Ruth Freitag can be persuaded to bring the work up to date in, say, 1990 when the present hullabaloo has died down.

Bruce Morton's bibliography will be less useful because he has deliberately omitted all non-English language source material. The listing is chronological but is backed up with an author and subject index. Morton only lists 1,301 items but does include many not covered by Freitag. Researchers will need both books.

Fire and Ice by Roberta J.M. Olson is an art historian's view of comets and has been published to accompany an exhibition of comets in art at the National Air and Space Museum in Washington, DC. The main attraction of Olson's book is the art reproductions. There are over a hundred pictures of comets, many in colour. Rackham, Moreau, Turner, Gérard, Blake, Daumier, van Stolk, Stech, Scott, Giotto — the list of artists is impressive. The text serves two purposes, first, to introduce the pictures and fit them into the context of the history of art, and secondly to weave in the development of man's fear and understanding of comets as shown in the paintings, engravings and illustrations of the day. All this is both entertaining and enlightening, but let me mention two minor quibbles. If someone has gone to the trouble to paint "Dr Halley" across the top of the Phillips portrait of our hero, why entitle it "Sir Edmund Halley" in the book? He wasn't knighted and he spelt his christian name *Edmond*. Also it must be remembered that Halley's comet was not *that* famous before 1705. Giotto might not have had to remember the comet of 1301; he might have popped out in December 1304 and painted the long-period comet that was in the sky then.

Peter Lancaster-Brown in *Halley and his Comet* has concentrated on the non-artistic historical aspects of the story, and creatively charts the transition of comets from astrological portents (in pre-Halleyan times) to their present status as unusual astronomical objects. The author has produced a most readable book, but one in which he has not resisted the temptation to fill in some of the gaps: lines such as "We might surmise without stretching credulity . . ." often end in rather too imaginative conjecture. If Newton was a latent homosexual, and Halley a philanderer, I would like chapter and verse.

All serious cometary historians should get a copy of *La Comète de Halley: Hier,*

Aujourd'hui, Demain by J. Alexandre and S. Debarbat. This is a small folder containing a collection of reproductions of Halley comet documents from the Paris Observatory. The Lubientetz maps and the drawings by Hevelius, Arnold, La Hire, J.-D. Cassini and Schwabe are fascinating, while Messier's hand-written account of the 1758 apparition is in the first division of historic documents.

Comet Fever by Donald Gropman has as its main theme the 1910 appearance of the comet and especially the dramatic events of the night of 18–19 May when the Earth swept through the comet's tail. We are told of sealed windows, absenteeism from home and work, mass prayer ses-



Richard Phillip's portrait of Halley circa 1721.

sions and, to cap it all, all the news fit to print about the infamous Oklahoma 19-year-old-naked-virgin-sacrifice-drama. Other chapters skim over past apparitions of Halley and aspects of modern cometary science, making in all a well-written, extremely readable book.

Fred L. Whipple is the doyen of comet scientists having worked extremely profitably in the field of cometary astronomy for 50 years. What he has to say about comets is certainly worth reading and his book, *The Mystery of Comets*, is superb. Too few scientists take the time to explain the complexities of their subject to a general audience. Whipple has left nothing out — if you want a synopsis of what is known about comets and, even more exciting, what is still beyond our present grasp, this is the book for you.

Another book that will be consulted for decades is *Halley's Comet in History*, edited by F.R. Stephenson and C.B.F. Walker, which emphasizes the Babylonian and Chinese records of the comet between 240 BC and AD 1531. The 13-page, line-by-line translation of a 164 BC Babylonian cuneiform tablet is rather strong meat for the casual browser, but if you need the historical information this is where to look. I wish more than one figure showing the comet's path across the sky had been included — the book would have been greatly enhanced by including a set of these for the relevant 23 apparitions.

Comet by Carl Sagan and Ann Druyan is a visual and intellectual feast, combining a host of specially commissioned colour illustrations with a text that, for the breadth of insight into the past, present

and future of cometary science, is way ahead of the rest of the field.

Standing back from all the razzamatazz surrounding this apparition, the book concentrates on the fundamentals: why have comets worried people, what have scientists and philosophers made of them, what is their origin, composition and fate? The Sagan's inquisitive enthusiasm is infectious. Sometimes they approach that danger zone where the gap between fact and speculation is overlooked, but they are soon forgiven. For me, comets will never be the same again. Cometary science can only blossom with such champions.

Finally, to the general books — those containing a bit of everything, history, comet lore and an observer's guide. Confronted with eight of them the reviewer is forced to list them in order of preference, a difficult task because they all have "curate's egg" properties. First past the post is Patrick Moore and John Mason's *The Return of Halley's Comet*. The maps are excellent, giving you a real chance of spotting the comet, and the review of comet science and decay is impressive. Make sure you get the new edition.

Close behind are Brian Harpur's *The Official Halley's Comet Book* and Mark Littmann and Donald K. Yeomans's *Comet Halley: Once in a Lifetime*. Both are superbly illustrated and overflowing with anecdotes. Harpur's review of cometary poetry and literature is masterly, while Littmann and Yeomans put across the scientific message superbly. With both, however, you will need additional help to find the comet.

In much the same category we find *A Comet Called Halley* by Ian Ridpath and Terence Murtagh. The text is excellent but some of the illustrations are a touch odd; surely the *Vega* spacecraft doesn't look like that! And it would be much easier to find the comet if it was shown with respect to compass points and elevation, and not just vaguely trundling across the constellations.

Occupying the uninspiring section where the errors proliferate and the illustrations become even more blurred we find four books by Richard Flaste *et al.*, John Tullius, Donald Tattersfield and Isaac Asimov. Reading these convinces me that there are many better ways of spending one's time and money.

The entire collection (which, incidentally, is by no means a comprehensive display of Halleyana) is listed in the table on the following page. But, to summarize: if you want to see Halley's comet read Francis Reddy; if you want to stay in your armchair with your feet up, read Brian Harpur, and Sagan and Druyan; if you want to know about comets in general you can do no better than Fred Whipple; and if you want to find something out for yourself get Ruth Freitag's bibliography. □

David W. Hughes is Senior Lecturer in Astronomy and Physics, Department of Physics, University of Sheffield, Sheffield S3 7RH.

Halley's comet in print — details of the publications

Title	Author/editor	Publisher	Pp.	Price	Rating
Seeing the comet					
Halley's Comet	Francis Reddy	AstroMedia, Milwaukee, WI/Pan, London	59	Pbk \$9.95, £7.95	*****
Halley's Comet Survival Kit	Tom Gaffney <i>et al.</i>	Perihelion Scientific, Paoli, PA	—	North America \$14.95; elsewhere \$20.25	**
The Night Sky Planisphere with Halley's Path	David Chandler	Sky Publishing, Cambridge, MA	—	\$5.95	**
Halley's Comet Finder	Ben Mayer	Perigee, New York	77	Pbk \$5.95	**
Mr. Halley's Comet: Everyone's Complete Guide to Seeing the Celestial Event	Editors of <i>Sky and Telescope</i>	Sky Publishing	32	\$2	**
Guides for teachers					
Comet Halley Returns: A Teachers' Guide 1985–1986	R.O. Chapman and R. Lynn Bondurant, Jr	NASA, EP 197	41	Pbk \$2	****
Halley's Comet Activities Manual	George S. Mumford	Sky Publishing	41	\$4	****
Halley's Comet: Teacher's Notes	Mary Ashby	Inner London Education Authority	53	£5.50	**
Specialist books					
Halley's Comet, a Bibliography	Ruth S. Freitag	Library of Congress, Washington, DC	555	\$26	*****
Halley's Comet, 1755–1984: A Bibliography	Bruce Morton	Greenwood Press, Westport, CT/Eurospan, London	—	\$35, £43.50	**
Fire and Ice: A History of Comets in Art	Roberta J.M. Olson	Walker, New York	144	Hbk \$24.95; pbk \$14.95	***
Halley and his Comet	Peter Lancaster-Brown	Blandford, Poole, Dorset/Sterling, New York	186	£9.95, \$12.95	**
La Comète de Halley: Hier, Aujourd'hui, Demain	J. Alexandre and S. Debarbat	Observatoire de Paris	—	FF 22	****
Comet Fever: A Popular History of Halley's Comet	Donald Gropman	Fireside (Simon & Schuster)	189	Pbk \$7.95	**
The Mystery of Comets	Fred L. Whipple	Smithsonian Institution Press/Cambridge University Press	276	Hbk \$24.95, £12.95; pbk \$12.50	****
Halley's Comet in History	F.R. Stephenson and C.B.F. Walker	British Museum Publications, London	64	Pbk £5.50	**
Comet	Carl Sagan and Ann Druyan	Random House/Michael Joseph	240	\$24.95, £12.95	****
General books					
The Return of Halley's Comet (new edition)	Patrick Moore and John Mason	Patrick Stephens, Wellingborough, Northants, UK/Warner, New York	128	Pbk £4.99, \$6.95	***
The Official Halley's Comet Book	Brian Harpur	Hodder & Stoughton, London/David & Charles, North Pomfret, VT	184	£8.95, \$15.95	**
Comet Halley: Once in a Lifetime	Mark Littman and Donald K. Yeomans	American Chemical Society, Washington DC	175	North America hbk \$19.95, pbk \$12.95; elsewhere hbk \$23.95, pbk \$15.95	**
A Comet Called Halley	Ian Ridpath and Terence Murtagh	Cambridge University Press	48	Pbk £2.95, \$4.95	**
The New York Times Guide to the Return of Halley's Comet	Richard Flaste <i>et al.</i>	Times Books, New York	244	Hbk \$16.95; pbk \$7.95	*
The Science Digest Book of Halley's Comet	John Tullius	Avon, New York	136	\$9.95	*
Halley's Comet	Donald Tattersfield	Basil Blackwell	164	£6.50, \$12.95	*
Asimov's Guide to Halley's Comet: The Awesome Story of Comets	Isaac Asimov	Walker, New York	118	\$12.95	*

Recent progress and future plans on the search for extraterrestrial intelligence

Michael D. Papagiannis

Department of Astronomy, Boston University, Boston, Massachusetts 02215, USA

The possibility of life in other parts of the Universe has long occupied the human mind, but actual searches only began in 1960 with Project OZMA conducted by Frank Drake. In the past 25 years, we have made impressive progress, and this new field has gained broad scientific recognition including the support of the US and the Soviet National Academies, and the endorsement of the International Astronomical Union.

PEOPLE have been fascinated by the possibility of intelligent life in other parts of the Universe for at least 2,500 yr (refs 1, 2). Around 400 BC, for example, the Greek philosopher Metrodorus of Chios wrote³: "It seems impossible in a large field to have only one shaft of wheat and in the infinite Universe only one living world." The question, however, remained a philosophical one until the middle of this century, when Cocconi and Morrison⁴ urged the scientific community to undertake radio searches in the vicinity of the 21-cm line of atomic hydrogen, the only radioastronomical line known at the time, for radio messages from other civilizations. They closed their paper with the statement: "The probability of success is difficult to estimate, but if we never search the chance of success is zero."

Independently, F. Drake, a young radioastronomer in the then embryonic staff of the National Radio Astronomy Observatory (NRAO) at Green Bank, West Virginia, was planning to undertake such a search with the 85-ft Tatel radiotelescope, the only one then existing at Green Bank. The search was undertaken in the spring of 1960, logging 200 h of observations of two Sun-like nearby stars, Epsilon Eridani (10.7 light yr) and Tau Ceti (11.9 light yr). He used a single-channel receiver and a new parametric amplifier, the best then available, which allowed him to reduce the system temperature from 1,500 K to ~350 K. The project was named OZMA after the Princess of the mythical kingdom of Oz. Drake⁵ reports the excitement of a false alarm they experienced when they first tuned their antenna to Epsilon Eridani. Such heart-stoppers have not been uncommon in projects associated with the search for extraterrestrial intelligence (SETI).

After a rather slow start in the 1960s (only three searches in the first 7 yr with less than 300 h of observations), SETI began to gain momentum in the 1970s, thanks to a meeting at the Byurakan Observatory⁶ in Soviet Armenia which was co-sponsored by the National Academies of both the Soviet Union and the United States, and a series of design studies⁷, workshops^{8,9} and conferences¹⁰, organized by the National Aeronautics and Space Administration (NASA) at its Ames Research Center. Finally in the 1980s it became established as an active branch of both the life sciences (exobiology) and of astronomy (bioastronomy). SETI was also included in the recommended astronomy projects for the 1980s by a special committee of the US National Academy of Sciences¹¹ and gained the endorsement of the International Astronomical Union (IAU), which at its 1982 triennial General Assembly established a new Commission¹² on this subject (IAU Commission 51—Search for Extraterrestrial Life). I was elected president, and N. Kardashev of the Soviet Union and F. Drake of the United States vice-presidents, all for the period 1982–85. The new commission grew rapidly to about 270 members and in 18–21 June 1984 organized in Boston the first IAU Symposium¹³ of this new branch of astronomy, which is rapidly becoming known as bioastronomy. During this symposium we celebrated 25 years since the historic paper by Cocconi and Morrison, while in the

spring of 1985 we celebrated the 25th anniversary of Project OZMA with a special meeting at NRAO in Green Bank which was attended by Drake and many of the pioneers and current workers in this field.

Search projects

In these 25 yr there have been about 50 search projects, most of them at radio frequencies but a few also in the optical and in the infrared. J. Tarter¹⁴ maintains the only active file of all these projects, which I have used as the basis for most of the descriptions that follow, augmented by personal information and data from papers by the experimenters included in ref. 13. So far we have amassed close to 120,000 h of observations using facilities in seven countries (United States, Soviet Union, Australia, Canada, France, Germany and Holland). Japan¹⁵ is also preparing to join the group with its new 45-m Nobeyama millimetre radiotelescope which is equipped with two acousto-optical spectrometers, one with 16,384 channels and a 2-GHz total bandwidth, and the other with 8,192 channels and a 160-MHz total bandwidth. They also have a 4,096 Fourier transform spectrometer with a frequency resolution of 150 Hz–10 kHz.

Almost all of the radio searches up to now have been conducted at selected 'magic frequencies', because from knowledge of science these radio lines would be known both to the transmitting and to the listening civilizations. This idea was advanced by Cocconi and Morrison⁴ who in essence said that if indeed 'they' want to make contact with us, 'they' will make it as easy as possible by transmitting at a frequency that is universally known. At the time, the 21-cm hydrogen line was the only one we knew, and though we now know radio lines of more than 65 atoms, molecules and radicals, the hydrogen line (*primus inter pares*) continues to dominate the searches. Unfortunately this reasoning has not paid off, but there was nothing else we could have done with the technology then available. In the next generation of searches, however, we plan to use multichannel spectrum analysers (MCSAs) with many millions of channels to search over a wide frequency range, rather than focusing only on certain magic frequencies.

Below we review the different searches that have been undertaken in the past 7 yr, or that are in progress. Following Tarter¹⁴, we will divide them into three general categories: Directed, Shared or Parasitic, and Dedicated, which are differentiated by the use they make of their observing facilities. For a description of searches before 1978 see ref. 16.

Directed searches

These searches use a major observatory for a specific SETI project of relatively short duration. Thus, in 1978, Sullivan *et al.*¹⁷ of the University of Washington proposed that an alternative search strategy would be to eavesdrop on nearby stars for radio signals leaking unintentionally into space from other advanced civilizations. As a test, Sullivan and Knowles¹⁸ used the large Arecibo antenna to observe the radio leakage of the

Earth in the 150–500-MHz range from its reflections from the Moon. Strong television stations and powerful military radars are the most prominent sources. As an example, the space surveillance radar of the US Navy in Archer City, Texas, which operates at 217 MHz emitting pulses of 1.4×10^{10} W into a bandwidth of only 0.1 Hz, could have been detected by a civilization with an Arecibo technology to a distance of 20 light yr. Knowles¹⁹ and Sullivan also did a limited eavesdropping search on a few nearby stars. The data were obtained with the Arecibo radiotelescope and the Mark I VLBI (very-long-baseline interferometry) system. They were analyzed for ultra-narrowband signals over a wide frequency range using a computer to create large Fourier transforms from tape-recorded one-bit time histories of the telescope voltage output. This technique, 'one-bit spectral analysis', is no substitute for a large MCSA, but can be used in parallel utilizing otherwise idle computer time. Although this test-run found no signals, it showed that eavesdropping on nearby stars is an alternative search technique that must be explored further.

Other directed searches in the United States include those of Tarter, Clark, Cuzzi, Duquet and Lesyna who used the Arecibo radiotelescope to observe 210 solar-type stars in a 4-MHz band around the magic frequencies of both HI and OH, and who also used the one-bit tape-recording technology and a CDC 7,600 computer to obtain a spectral resolution of 5.5 Hz. Using the 14-m millimetre radiotelescope of the University of Massachusetts, Lord and O'Dea looked for powerful beacons at the 115-GHz (2.6-mm) line of CO along the north rotational axis of the Galaxy, a potentially 'magic location' for beacons operated by supercivilizations.

On the other side of the Atlantic, F. Biraud of the Observatory of Meudon and J. Tarter of Berkeley and NASA-Ames have been using the large (40 × 240 m) radiotelescope at Nançay, France, since 1981, and an eight-level 1,024-channel autocorrelator with a resolution of 50 Hz per channel to observe 300 solar-type stars at four OH frequencies. This project is complementing previous work done with the Arecibo and the 300-ft NRAO radiotelescopes, where low declination sources had not been observed. In Holland, Shostak of the Kapteyn Astronomical Institute and Tarter used in Project SIGNAL the Westerbork array to look with high angular resolution (to screen out noise) for a pulsed beacon, with a repetition rate between 40 s and 1 h, at the centre of the Galaxy, a magic location for a supercivilization.

P. Horowitz of Harvard, while on sabbatical at Stanford, built together with Teague, Linscott, Chen and Backus, an impressively compact MCSA, which they named 'Suitcase SETI'. It had 65,536 (2^{16}) channels for each of the two circular polarizations and a resolution of 0.03 Hz per channel. They used it in Arecibo to observe 250 solar-type stars at the hydrogen line and at twice the hydrogen frequency. They were looking for an ultra-narrowband signal, which is the most economic means for interstellar communications. The Drake-Helou spreading²⁰ due to multiple scattering from inhomogeneities of the interstellar medium, limits the narrowness of the signal to a bandwidth of 0.01–0.1 Hz. This initial experiment later led to Project SENTINEL. Another interesting search was undertaken in 1983 by R. Freitas and F. Valdez who used the Hat Creek radio observatory to conduct a search at the 1,516-MHz line of tritium. Since tritium has a half-life of only 12.5 yr, if found in the vicinity of normal stars it could only have an artificial origin, possibly the by-product of a huge nuclear fusion plant. The same workers had earlier used the 30-inch telescope of the University of California and the 24-inch telescope of the Kitt Peak National Observatory (KPNO) to conduct an optical search for artefacts and probes placed in our Solar System by extrasolar civilizations. They looked at the magic locations of the L4 and L5 Lagrange points of the Earth-Moon and Earth-Sun systems, which are the most stable regions in space to observe the Earth.

Other non-radio search projects include those of V. Shvartsman and his colleagues of the Special Astrophysical Observatory of the Soviet Union, who used the 6-m optical telescope at

Zelenchuskaya, USSR, and a device called MANIA which can detect very short pulses (3×10^{-7} – 3×10^2 s) from optical lasers, to look at 21 stellar objects with unusual spectral properties, hypothesizing that they may be examples of supercivilizations that could be using laser signals for interstellar communications. In a related project, a laser heterodyne receiver operating at 30 THz (10 μ m) is now being built by H. Betz of the University of California-Berkeley that will have 1,000 channels each 1 MHz wide, to undertake an infrared search for extraterrestrial laser signals which in certain circumstances might be used for interstellar communications. Finally, I²¹ am preparing to undertake a search for large artificial objects (such as space colonies and materials processing plants) in our Solar System and especially in the asteroid belt, which is an ideal source of raw materials, to test the theory of galactic colonization, that is, the possibility that the entire Galaxy might already have been colonized by more advanced civilizations. I will be using the Infrared Astronomy Satellite (IRAS) data bank of infrared sources in our Solar System, which is now being prepared at the Jet Propulsion Laboratory by the IRAS Asteroid Workshop²², to look for objects with unusual infrared signatures in our Solar System.

Shared or parasitic searches

These searches either re-analyse for ETI signals old, archived data that had been obtained for other purposes, or they share in a parasitic or piggyback fashion the data being obtained by a radiotelescope for an unrelated project and process them for ETI signals in real time. Thus by re-analysing many of the 21-cm 'noisy' sky maps that had been obtained with the Westerbork array and had been kept by the Dutch astronomers, F. Israel of the Netherlands working first with DeRuiter and then with Tarter, looked for strong sources in positions that coincided with those of known stars. It is hoped that a similar process will be used in the future with maps obtained with the VLA (Very Large Array). Meanwhile, Cohen *et al.*²³ re-analysed their own surveys of globular clusters, looking for narrowband signals at frequencies associated with OH and H₂O masers that might be tapped by supercivilizations to produce strong signals in certain directions. M. Damashek is re-analysing for ETI signals 700 h of data that he had obtained with the NRAO 300-ft antenna during an extensive pulsar search at 390 MHz over most of the sky.

In 1980, S. Bowyer and D. Werthimer of the University of California-Berkeley and their collaborators^{24,25}, built an automated 100-channel spectrum analyser to siphon through it, from the intermediate frequency (IF) stage, data obtained by a telescope for an unrelated project. This parasitic or piggyback SETI device was named SERENDIP and was used with the Hat Creek and Golstone antennas. It was recently upgraded to SERENDIP II, which has a 65,536-channel fast Fourier processor with a resolution of 2 Hz per channel. It will be used to search for narrowband peaks above a 6σ threshold over the entire 30-MHz IF band of the radiotelescope searching through increments of 100 kHz in 10 s or less. Data of unidentified peaks are recorded for further investigation. The system will operate unattended on a 24-h basis with practically any radiotelescope.

Many other parasitic searches have been carried out around the world. Thus while searching for impulsive radio events at 5 GHz with the 64-m Parkes radiotelescope in Australia, T. Cole and R. Ekers also checked a few G and K main-sequence stars for ETI signals of short duration. In a 4-day run they detected only one such signal in the direction of 82-Eridani which lasted only 1 ms. Efforts to re-acquire it were unsuccessful. During a routine search for pulsars with the large 100-m Bonn (Effelsberg) antenna, R. Wielebinski and J. Seiradakis of the Max Planck Institute for Radio Astronomy, looked for ETI pulsed signals at the hydrogen line with repetition periods between 0.3 and 1.5 s. And during a study of the galactic magnetic field with the 46-m Algonquin radiotelescope of Canada at 10.6 GHz (2.8 cm), Vallee²⁶ and Simard-Normandin of the Herzberg Institute also looked for highly (>70%) linearly polarized signals, which

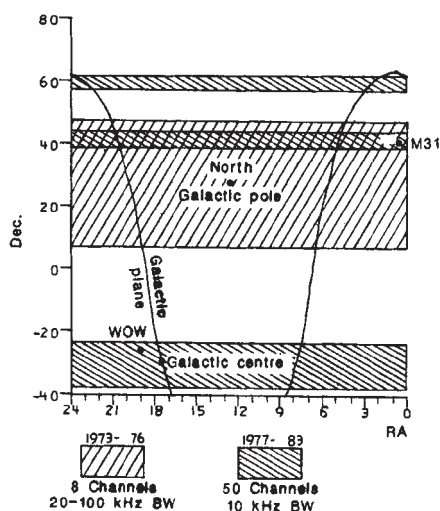


Fig. 1 The regions of the sky that have been surveyed so far by the Ohio SETI Program³².

could be strongly suggestive of an artificial origin, near the centre of the Galaxy where a supercivilization could be operating an intergalactic relay station. Taking advantage of the fact that the 64-m antenna of NASA's Deep Space Station in Tidbinbilla, Australia, was stowed for several months for structural repairs, Gulkis, Kuiper, Olsen, Jauncey and Peters of the Jet Propulsion Laboratory used it for a SETI sky survey at 8 and 22 GHz with a 256-channel spectrum analyser. The elevation of the telescope was occasionally changed to scan a different band of the sky passing over it. Finally, the 1.5-m Mt Lemon optical telescope was used by F. Witteborn of NASA-Ames to further study 20 stars that in a routine infrared study had been found to be faint in the optical for their spectral type, but strong in the infrared. This made them potential candidates for Dyson Spheres, that is, the products of astroengineering activity by supercivilizations which could use all the planetary matter of their solar system to build a huge shell around their central star to tap a large fraction of its energy. Nothing was discovered that could not be explained by natural processes.

Turning to the Soviet Union, Slyph²⁷ and Kardashev²⁸ have also been looking in recent years for Dyson Spheres and other superstructures of astroengineering activity, searching primarily through the IRAS data for stellar sources with an unusual infrared signature. An all-sky search for microwave background fluctuations is under consideration by the Soviet workers to back up the IRAS data. Unfortunately during the past seven years there have been no radio searches in the Soviet Union, where there was considerable SETI, or CETI (Communication with ETI) as Soviet workers prefer to call it, activity in the 1960s and 1970s^{29,30} under the leadership of V. S. Troitskii and N. S. Kardashev. In the Proceedings of the Montreal meeting in 1979³¹ and at the SETI '81 meeting in Tallinn Estonia, Soviet workers had reported that they were preparing to start building in Gorki under the direction of Troitskii an array of ~100 1-m parabolic antennas (Project OBZOR) for SETI work at 21 cm. During my visit to Moscow in August 1984, I was told that this project had been abandoned, primarily because of the long illness of Troitskii—a pioneer and an international authority in this field. The emphasis in the Soviet Union is now geared towards the detection of astroengineering structures by supercivilizations, but I feel that under the leadership of Kardashev and Slyph, they will regain their high-ranking position in the radio searches, especially with the use for SETI of the new 65-mm Samarkand millimetre radiotelescope now approaching completion.

Dedicated searches

Dedicated searches are performed by radio observatories that have become SETI dedicated facilities and conduct searches on a continuous basis. There are two such projects now in operation,

which account for ~80% of the observing hours accumulated thus far. The oldest of the two is the Ohio SETI Program, having been in operation since 1973 under the direction of J. Kraus and R. Dixon³². It uses the meridian-transit radiotelescope of the Ohio State University which has a collecting area of 2,200 m², equivalent to a parabolic dish 53 m (175 ft) in diameter, and a beam pattern at 21 cm of 8 arc min in RA and 40 arc min in dec. The search is conducted at the hydrogen line, Doppler shifted in the rest frame of the galactic centre, with the help of a 50-channel filter bank with a resolution of 10 kHz per channel. During the first 3 yr (1973–76), however, the system was less sophisticated and operated with only eight channels. Figure 1 shows the portions of the accessible sky (–36 to +63 in declination) that have been covered thus far by the old and the new system. The sky band 40–43 dec was chosen because it includes the Andromeda Galaxy (M31), our large neighbouring galaxy.

The 'WOW' signal of the Ohio SETI Program (Fig. 1) was recorded in 1977 at –27 dec. It earned its name from the word the telescope operator penned next to this peak on the computer printout of the telescope, and in spite of repeated efforts by the Ohio group and others it has never been re-acquired. The data obtained are archived and analysed for statistical effects with respect to celestial coordinates. Several improvements are being planned for the future, including expanding the frequency range to cover the entire 'water hole' (1,400–1,750 MHz) which is bracketed by the hydrogen and the OH lines, that together form water. The Ohio SETI Program has been sustained with modest support from NASA, the tireless efforts of Kraus and Dixon, and the unselfish dedication of many enthusiastic volunteers.

Horowitz³³ of Harvard University and his collaborators have been operating since March 1983 the other dedicated facility at the Oak Ridge Harvard-Smithsonian Observatory near Boston. It is called Project SENTINEL and is supported by the Planetary Society, a private organization headed by C. Sagan and B. Murray. It is an outgrowth of project Suitcase SETI mentioned earlier, and uses the same two 65,536-channel spectrum analysers, with a resolution of 0.03 Hz per channel, with the 84-ft antenna of the Oak Ridge Observatory. With the telescope set at a particular declination, the search sweeps a 0.5° band around the sky, covering the entire available sky (about 80% of total) in ~1 yr. During their 2 yr of operation (1983–85) they have covered all of the accessible sky, first at the hydrogen line at 1,420.40575 MHz and then at one of the four OH lines at 1,667.3590 MHz. A potential source stays in the antenna beam for about 2.5 min. The system searches automatically for large peaks in any of its channels and archives anything suspicious. Because of the expected swept-Doppler signature produced by the Earth's rotation, the system is very good in rejecting radio interference, though they too had two false alarms of almost 50σ when they got the Sun into their beam.

The problem with this search is its very narrow (2 kHz) total bandwidth around the hydrogen line. This necessitates a continuous correction for all the Doppler shifts due to the motions of the Earth relative to the direction of the antenna, and assumes that the transmitting civilization will also correct for all Doppler shifts, including the relative motion of our two stars (the radial component of the peculiar motion) which can be of the order of 100 kHz, that is 50 times larger than the 2-kHz total bandwidth of the system. It also requires that they beam their signals specifically to our Solar System. To overcome these restrictions, Horowitz and his colleagues³⁴ have built a brand new MCSA with 8.4×10^6 channels and a frequency resolution of 0.05 Hz per channel, which is giving them a total bandwidth of 420 kHz, enough to account for practically all of the Doppler effects. In the 2.5 min that a source stays in their antenna beam the system will make six and possibly eight measurements checking the left-hand and the right-hand circular polarization at three or four distinct frequencies of the same line resulting from Doppler effects in three or four different rest frames (the local standard of rest of our Solar System and possibly their heliocentric frame, the reference frame of the galactic centre, and the cosmic black-body rest frame in which the 3 K background radiation is

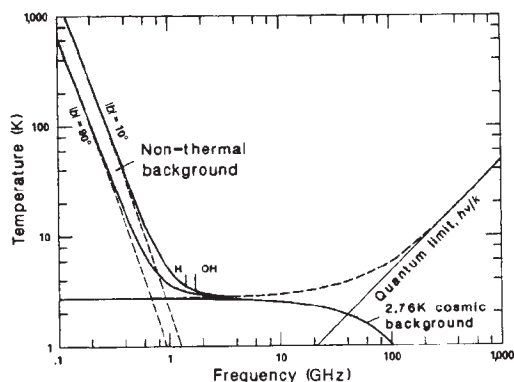


Fig. 2 The microwave window of interstellar space³⁸.

isotropic). This new system, which has been named Project META, went into operation on September 29, 1985.

Technological progress and NASA's Program

The search for ETI radio signals is a multidimensional problem which because of its complexity is often called the 'cosmic haystack'. The dimensions of the search space include: location of sources; transmitting frequency; signal strength; bandwidth; polarization; signal modulation; and on-off periods. Initially investigators in the United States had hoped that ETI would be transmitting at the hydrogen line to make contact easy. But extensive, although not exhaustive, searches by many investigators at hydrogen and a few other magic frequencies have not produced any positive results, which indicates that the search is not going to be an easy task. With the improvement in our technology, NASA is now preparing to embark on the next generation of radio searches, that is a systematic search over a wide frequency range (rather than the old magic frequencies) in the microwave window (1–10 GHz) of the Earth. This is the free-space microwave window seen on Fig. 2, cut off for ground observations around 10 GHz by the water vapour in the Earth's atmosphere. This frequency range is relatively free of natural radio noise, and is therefore best suited for interstellar communications. Embedded deep in the microwave window is the 'water hole' (1.4–1.7 GHz), probably the most attractive frequency range for interstellar radio communications. A point of great concern, however, is the rapidly growing use of this frequency range for other purposes, which are bound to interfere with future radio searches.

It is generally believed that radio signals will be narrowband to reduce peak power. To have a good signal-to-noise ratio we need receivers with very narrow bandwidths, and since we need to explore a wide frequency range we need a large number of narrowband channels. This function is fulfilled by a MCSA³⁵, an electronic high-resolution radio spectrometer, which in the past 25 yr has evolved from a few channels to 65,000 channels and we are soon to have MCSAs with 8×10^6 channels and possibly more. We have already mentioned the ultra-narrowband MCSA constructed by Horowitz for his Project META. NASA is also planning an 8.25×10^6 -channel MCSA, but with a much wider total bandwidth, which is now being developed by a group at Stanford University headed by Peterson³⁶ and Linscott. It will consist of two cascaded banks of digital filters followed by multiplexed microprocessors executing digital Fourier transforms. The complete MCSA will have 112 units, each with 73,728 channels for a total of 8,257,536 channels. The highest frequency resolution will be 1 Hz per channel, giving a total bandwidth of about 8 MHz. It will also be able to analyse simultaneously the incoming 8-MHz frequency band into channels of 32, 1,024 and 73,728 Hz. In all cases, it will examine both the left-hand and the right-hand circular polarizations.

The MCSA performs a complex Fourier transform that yields the real and imaginary amplitude of the signal, which are then

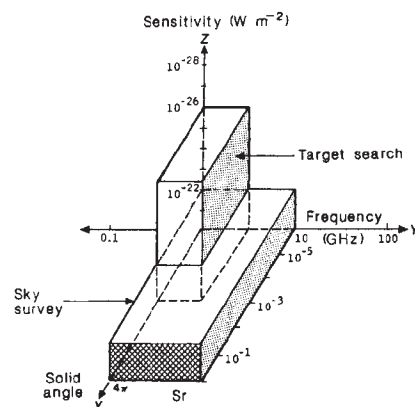


Fig. 3 The two 'volumes' of the SETI Search Space, that will be searched by the Sky survey and the Targeted search, the two complementary components of the NASA SETI Program⁴⁰.

squared and added to give the actual power for each bin (channel). If it exceeds a predetermined threshold, chosen by the desired signal-to-noise ratio, the signal is flagged for further tests. Sophisticated signal detection algorithms^{35–37} are now being developed by NASA. Considerable effort is being made to achieve online processing of the data, which is not an easy task given the huge volume of incoming data with an 8×10^6 -channel MCSA. Emphasis is also placed on the ability to detect pulsed signals and signals with a Doppler frequency drift³⁷. The 8×10^6 channel MCSA could be ready by 1988, but more realistically the whole system, including all the signal processing instrumentation, will probably become operational in the early 1990s. A prototype unit with 73,728 channels, together with several of the new signal recognition algorithms, is now undergoing tests at the Goldstone antenna. In a recent test, they were able to pick up the very weak signal (1 W) beamed by Pioneer 10 towards the Earth from a distance of ~ 35 AU ($\sim 5,250 \times 10^6$ km). They were also able to see clearly the frequency drifts imposed on this monochromatic signal by the rotation of the Earth and the relative motion of the Earth and the spacecraft.

The NASA SETI Program³⁸, which has been in the study and development stage for the past 15 yr, is now headed by B. Oliver, one of the pioneers of SETI⁷, and includes about 15 well-known scientists at the three California institutions (NASA-Ames, Jet Propulsion Laboratory, and Stanford University) that share responsibility for this programme. The NASA SETI Program will be a bimodal search consisting of two components: the Targeted search³⁹, which will emphasize sensitivity to weak signals by concentrating only on a number of discrete sources, and the Sky survey^{40–42}, which will emphasize sky coverage by scanning the entire sky. Figure 3 shows the relative volumes of the search space that will be investigated by these two complementary searches.

The Targeted search will focus on 800–1,000 specific targets, including the 773 F, G and K Sun-like stars up to a distance of 25 pc (81.5 light yr) included in the Royal Greenwich Observatory Catalogue, and a variety of other targets including stars with peculiar spectra and galaxies. It will have a peak spectral resolution of 1 Hz and will concentrate in the frequency range surrounding the water hole. The sensitivity (Φ) of radio observations, usually given in W m^{-2} , is related to the demanded signal-to-noise ratio (a), the system temperature (T), the collecting area of the antenna (A) which is proportional to the square of the diameter (D) of the antenna, the resolution bandwidth (b), and the integration time (τ), by the expression:

$$\Phi = a \frac{kT}{A} \left(\frac{b}{\tau} \right)^{1/2} \quad (1)$$

Since for given Φ , T and b , τ is proportional to D^4 , to achieve the highest possible sensitivity in a reasonable time, the targeted

search will use the largest possible antennas including the 305-m Arecibo, the 53-m (equivalent) Ohio State, the 64-m Goldstone (California) and Tidbinbilla (Australia) antennas of NASA's Deep Space Network, the 91-m NRAO, and possibly the 100-m Bonn (Germany) antenna. It is interesting to note that when the 1,000-ft Arecibo antenna with a system temperature of 35 K is compared with Drake's 85-ft NRAO antenna with a 350 K system temperature, Arecibo would be able to do the 200 h of Project OZMA in just a fraction of a second.

An approximate time calculation for the Targeted search yields $1,000 \text{ targets} \times 1,000 \text{ s per target per 8-MHz frequency band} \times 100 \text{ frequency bands}$ to cover the range 1.2–2.0 GHz around the water hole $= 10^8 \text{ s} = 3.3 \text{ yr}$, which, including losses of time for moving the antenna from target to target, antenna maintenance, transportation of the MCSA from station to station, and so on, amounts to $\sim 5 \text{ yr}$ of telescope availability. The Sky survey will use several of the smaller, 34-m radio-telescopes in NASA's Deep Space Network. It will have a lower spectral resolution of 32 Hz per channel, but will cover a wider frequency range (1–10 GHz). It will survey the entire sky ($4\pi \text{ sr}$) with a half-power beamwidth (HPBW) of $\sim \lambda/D$ and therefore will, in essence, examine $4\pi/(\lambda/D)^2 \sim 10^6$ locations. Spending from 0.3 to 3 s per location (to scan through a HPBW) and per frequency band of about 250 MHz, and needing ~ 36 such bands to cover the 1–10 GHz range, yields again a total of $\sim 10^8 \text{ s}$, that is also $\sim 5 \text{ yr}$ of telescope availability. Thus the entire NASA SETI Program (Targeted search and Sky survey) will require $\sim 10 \text{ yr}$ telescope availability. Using 5–10 different radiotelescopes at ~ 10 –20% of their time, and assuming the availability of several MCSA units, the whole programme will take optimistically $\sim 5 \text{ yr}$, but more realistically probably close to 10 yr. Hence assuming a starting date around 1990, we can expect it to be completed around the year 2000, having enhanced immensely the volume of the search space that we will have explored. It is also possible that a high spectral resolution search, such as this, over the entire sky and over a wide frequency range, might produce also some unexpected astronomical discoveries.

Rationale behind the search strategy

Sparked off by a critical paper by Hart⁴³ in 1975, there were many heated debates⁴⁴ and counter-arguments⁴⁵ in the late 1970s and early 1980s on the most advisable search strategy and even on the advisability of undertaking any searches at all. The key points to these debates were: (1) the possibility of interstellar travel and of galactic colonization, and (2) the lack of any scientifically verifiable evidence of visits to Earth by extraterrestrials, which is often referred to as the 'Fermi Paradox' or the 'Great Silence'⁴⁶. After many debates, we began to realize that none of us can claim to know how civilizations far more advanced than ours are likely to behave and act. The conclusion, therefore, which was stated by several participants^{46–50} at the recent IAU Symposium¹³ can be summarized as follows: debates, useful as they might be to sharpen our understanding of the issues, will never answer our questions. The answers can come only through scientific searches, which must be pursued vigorously. The search strategy, however, ought to be broad and flexible enough to allow, besides the mainstream searches such as the NASA SETI Program, the parallel experimental testing of different other theoretical possibilities. With such a broader search strategy we are likely to keep many more good people active in this young field, and at the same time increase our chances of success by not neglecting any of the possible alternatives.

From a comparison of Project OZMA and Arecibo with an MCSA, it is clear that we have made colossal technological progress in less than three decades. A question often asked is why we do not postpone our searches until our technology becomes more effective. I believe the answer has two parts. The first is that we do not know in advance the level of technology that would be sufficient. It would have been a grave mistake to have asked Wilbur and Orville Wright to wait for the discovery of the jet engine before trying to fly. They succeeded with far

less, and on 17 December 1903 they opened the doors of the new field of aviation. I am sure that even if Drake could have known that 25 yr later Horowitz would have had an 8×10^6 MCSA for a search around the hydrogen line, he still would have gone ahead with his Project OZMA, and rightfully so. The second reason is that technology is like a ladder that we must climb one rung at a time, starting from the lowest ones. But as we start climbing, the horizons broaden and many new technological developments materialize, such as the jet engine in aviation and the 8×10^6 MCSA in SETI, with benefits also for many other fields.

After millennia of thinking and philosophizing about the 'plurality of worlds', we have finally entered the experimental era. We have already used space probes to search for primitive life in our Solar System, and we are now searching with radio, optical and infrared telescopes for other advanced civilizations in the Galaxy. It is a special privilege to live in the era that tries to answer experimentally profound old questions about the prevalence of life, and especially of life with intelligence, in the Universe. With the NASA SETI Program as the central force, and the many other parallel special searches now in progress or planned for the future, we can expect that in the next 10–20 years we will know much more about the presence of other advanced civilizations in our Galaxy. If we were to find them, this would be the greatest discovery of all time. But even if after concerted efforts we were to conclude that we must be one of very few if not the only advanced civilization in the Galaxy, this too would be an important discovery, because knowing how rare our civilization is among the hundreds of billions of stars in our Galaxy would hopefully make us realize how cosmically important it is to preserve it.

I thank the many contributors to the *Proceedings of the IAU Symposium 112*¹³ whose works helped me in the preparation of this review. I also thank my colleagues, especially J. Billingham, D. DeVincenzi, R. Dixon, F. Drake, S. Gulkis, H. Hirabayashi, P. Horowitz, J. Jugaku, N. Kardashev, M. Klein, B. Oliver, C. Sagan, C. Seeger, V. Slysh, W. Sullivan, J. Tarter and V. Troitskii, for useful discussions and comments on the first draft of this article. This work was supported by NASA grant NAGW-597.

1. Bell, T. E. *Griffith Observer* 42, No. 8, 2 (1978).
2. Dick, S. J. *Plurality of Worlds* (Cambridge University Press, 1982).
3. Diels, H. *Die Fragmente der Vorsokratiker*, 141 (Weidmannsche Buchhandlung, Berlin, 1922).
4. Cocconi, G. & Morrison, P. *Nature* 184, 884 (1959).
5. Drake, F. D. *Cosmic Search* 1, No. 1, 11 (1979).
6. Sagan, C. (ed.) *Communications with Extraterrestrial Intelligence* (MIT Press, Cambridge, 1973).
7. Oliver, B. M. & Billingham, J. (eds) *Project Cyclops* (NASA CR11445, 1973).
8. Ponnampetuma, C. & Cameron, A. G. W. (eds) *Interstellar Communications: Scientific Perspectives* (Houghton Mifflin, Boston, 1974).
9. Morrison, P., Billingham, J. & Wolfe, J. (eds) *The Search for Extraterrestrial Intelligence SETI* (NASA SP-419, 1977).
10. Billingham, J. (ed.) *Life in the Universe* (MIT Press, Cambridge, 1981).
11. Field, G. B. (ed.) *Astronomy and Astrophysics for the 1980s* (National Academy Press, Washington, 1982).
12. Papagiannis, M. D. *J.B.I.S.* 36, 305 (1983).
13. Papagiannis, M. D. (ed.) *IAU Symp. 112* (Reidel, Dordrecht, 1985).
14. Tarter, J. *IAU Symp. 112*, 271–290 (1985).
15. Hirabayashi, H. *IAU Symp. 112*, 425–434 (1985).
16. Zuckerman, B. & Tarter, J. in *Strategies for the Search for Life in the Universe* (ed. Papagiannis, M.D.) 93–106 (Reidel, Dordrecht, 1980).
17. Sullivan, W. T. III, Brown, S. & Wetherill, C. *Science* 199, 377 (1978).
18. Sullivan, W. T. III & Knowles, S. H. in *IAU Symp. 112*, 327–334 (1985).
19. Knowles, S. H. *IAU Symp. 112*, 335–340 (1985).
20. Drake, F. D. & Helou, G. *Rep. 76 NAIC* (Cornell University, Ithaca 1977).
21. Papagiannis, M. D. *IAU Symp. 112*, 505–511 (1985).
22. Tedesco, E. F. (ed.) *IRAS Asteroid Workshop No. 4* (JPL D-2176, 1985).
23. Cohen, N., Malkan, M. & Dickey, J. *Icarus*, 41, 198 (1980).
24. Bowyer, S., Zeitlin, G. M., Tarter, J., Lampton, M. & Welch, W. *Icarus*, 53, 147 (1983).
25. Werthimer, D., Tarter, J. & Boyer, S. *IAU Symp. 112*, 421–424 (1985).
26. Vallee, J. P. *IAU Symp. 112*, 321–326 (1985).
27. Slysh, V. I. *IAU Symp. 112*, 315–320 (1985).
28. Kardashev, N. S. *IAU Symp. 112*, 497–504 (1985).
29. Troitskii, V. S. *et al. Soviet Astr.* 18, 669 (1975).
30. Troitskii, V. S. *et al. Acta astronaut.* 6, 81 (1979).
31. Papagiannis, M. D. in *Strategies for the Search for Life in the Universe* (ed. Papagiannis, M. D.) 3–12 (Reidel, Dordrecht, 1980).
32. Dixon, R. S. *IAU Symp. 112*, 305–314 (1985).
33. Horowitz, P. & Forster, J. *IAU Symp. 112*, 291–304 (1985).
34. Horowitz, P., Forster, J. & Linscott, I. R. in *IAU Symp. 112*, 361–372 (1985).
35. Oliver, B. M. *IAU Symp. 112*, 343–350 (1985).

36. Peterson, A. M., Chen, K. S. & Linscott, I. R. in *IAU Symp.* 112, 373-384 (1985).
37. Cullers, K. D. *IAU Symp.* 112, 385-390 (1985).
38. Drake, F., Wolfe, J. H. & Seeger, C. L. (eds) *SETI Science Working Group Report* (NASA Tech. Pap. 2244, 1983).
39. Seeger, C. L. & Wolfe, J. H. *IAU Symp.* 112, 391-398 (1985).
40. Klein, M. J. & Gulkis, S. *IAU Symp.* 112, 397-404 (1985).
41. Olsen, E. T., Lokshin, A. & Gulkis, S. *IAU Symp.* 112, 405-410 (1985).
42. Gulkis, S. *IAU Symp.* 112, 411-418 (1985).
43. Hart, M. H. *Q. Jl. R. astr. Soc.* 16, 128 (1975).
44. Papagiannis, M. D. (ed.) *Strategies for the Search for Life in the Universe*, 13-76 (Reidel, Dordrecht, 1980).
45. Hart, M. H. & Zuckerman, B. (eds) *Extraterrestrials, Where Are They?* (Pergamon, Oxford, 1982).
46. Brim, G. D. *Q. Jl. R. astr. Soc.* 24, 283 (1983).
47. Morrison, P. *IAU Symp.* 112, 13-20 (1985).
48. Papagiannis, M. D. *IAU Symp.* 112, 505-512 and 543-546 (1985).
49. Seeger, C. L. *IAU Symp.* 112, 487-492 (1985).
50. Smith, H. J. *IAU Symp.* 112, 547-552 (1985).

ARTICLES

Intensive measurements of turbulence and shear in the equatorial undercurrent

M. C. Gregg*, H. Peters*, J. C. Wesson*, N. S. Oakey† & T. J. Shay*

* Applied Physics Laboratory and School of Oceanography, College of Ocean and Fishery Sciences, University of Washington, Seattle, Washington 98195, USA

† Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada

Average profiles of turbulent dissipation rates provide tests of parameterizations used in numerical models. A strong diurnal mixing cycle is found in the well-stratified high-shear zone above the velocity maximum, and thermohaline intrusions are shown to be important mixing agents below the velocity maximum.

INTENSE and persistent mixing occurs in the equatorial undercurrents of the Atlantic and the Pacific oceans. Determining and parameterizing the average turbulent fluxes is necessary because these fluxes may provide the primary dynamical balance to the pressure gradient driving the undercurrent. Turbulent transport in the undercurrent is particularly important in the central Pacific, where it is a major factor controlling the tongue of cold water extending westward from South America¹. Temperature changes of the cold tongue affect the air-sea interactions that lead to anomalous weather conditions over North America. Therefore, improving the measurement and parameterization of the vertical turbulent fluxes in the undercurrent is a major goal of the Tropic Heat Program established to study the cold tongue¹. The undercurrent also provides an excellent natural laboratory for studying stratified turbulence.

Although the amount of previous turbulence data from the undercurrent is small, microstructure profiles taken at various locations and times, with different instruments, show similar patterns and intensities²⁻⁵. Earlier measurements with towed bodies report levels at the core of the undercurrent that are several decades higher than those observed with profilers^{6,7}, leading some to conclude that the towed data are badly contaminated by vibration and sensor difficulties^{2,5}. Others maintain that large intermittence at the velocity maximum produces the disparity⁸.

Between 25 and 30 November 1984, during the Tropic Heat field measurements, continuous profiling of turbulence was done in the upper 150 m at 0° N, 139°50' W with the advanced microstructure profiler (AMP)⁹. In 4.5 days, 385 drops were obtained, providing enough data to form accurate ensemble average profiles and to examine short-term fluctuations. Also, eight shear measurements were obtained with the expendable current profiler (XCP)¹⁰.

Background conditions

The average stratification and shear are typical of conditions near 0° N, 140° W. The mean stability frequency, $\bar{N}$, is low near the surface but increases rapidly, reaching a maximum of 0.02 s^{-1} near a pressure (p) of 1.0 MPa (corresponding to a depth of 100 m). Close to the surface the stratification is controlled by temperature, but salinity is important below 0.65 MPa. Because the XCP measures relative velocities, the average speed between 5.0 and 5.5 MPa is set to zero, in general agreement with mooring observations. Useful signals from the XCP begin at 0.3 MPa;

the transport at shallower depths is to the west, as observed from the ship drift. The zonal flow below 0.3 MPa is eastward (zonal velocity $\bar{U} > 0$ in Fig. 1), reaching 1.6 m s^{-1} near 1.05 MPa, which is close to the depth of the salinity maximum. Owing to the proximity of the magnetic equator the north component of the XCP signal is not useful. Nevertheless, because of the dominance of the eastward flow, good estimates of shear are obtained from $\partial \bar{U} / \partial z$ (where z is the depth). Values are calculated from the slopes of linear fits to the raw velocity components. The interval of each fit is 0.2 MPa, and a fit is done every 0.05 MPa. Shear decreases with depth and has a local minimum at the velocity maximum. Because this occurs close to $\bar{N}_{\text{max}}$, the Richardson number, $Ri = \bar{N}^2 (\partial \bar{U} / \partial z)^{-2}$, has a sharp maximum. Above 0.5 MPa, $Ri < 0.25$; between 0.5 and 0.9 MPa, it is about 0.35 (Fig. 2). Based on fluctuations in the eight shear profiles used to form the average, we estimate typical confidence limits to be (0.0-0.15) for $Ri = 0.07$ and (0.2-0.5) for $Ri = 0.3$. Vertical smearing, due to displacements by internal waves and the internal tide, makes the velocity maximum appear thicker and reduces the maximum value. $Ri < Ri_{\text{cr}} = 0.25$ is a necessary but not sufficient condition for instability of laminar flows¹¹.

Average mixing rates above velocity maximum

Centimetre-scale gradients of horizontal velocity and temperature measured with AMP are used to compute the variances of the gradients over successive 0.005 MPa (0.5-m) sections. Scaling yields $\epsilon = 7.5 \nu \langle (\partial U' / \partial z)^2 \rangle$ and $\chi = 6 \kappa_T \langle (\partial T' / \partial z)^2 \rangle$. ϵ is the rate of viscous dissipation of turbulent kinetic energy, and χ is the rate of diffusive smoothing of turbulent temperature fluctuations. (ν is the kinematic viscosity and κ_T is the thermal diffusivity.) In both cases, isotropy of the microscale gradients is assumed.

Because of the large number of records taken in a short time, average profiles of ϵ and χ can be formed using all data in each 0.005 MPa pressure bin (Fig. 3). Comparison with laboratory results provides a measure of the relative intensity of $\bar{\epsilon}$. Studies of decaying grid turbulence in unsheared flows¹² show that when $\epsilon > \epsilon_{\text{cr}} = 24.5 \nu N^2$, the turbulence will be sufficiently intense to produce a negative buoyancy flux, that is, $J_b = -g \rho^{-1} w' \rho' < 0$ where g is the gravitational acceleration, ρ' is the density and w' is the vertical velocity. A negative buoyancy flux corresponds to an upward mass flux, which increases the mean potential energy. Preliminary results from sheared flows¹³ show that the

Fig. 1 Average profiles of the stratification and velocity fields; θ is potential temperature, s is salinity (in concentration units), σ_θ is potential density, p is pressure, and U is the zonal velocity (positive east). These plots are formed from averages of 385 AMP profiles and eight XCP profiles. Owing to vertical displacement by internal waves and the internal tide, averaging smooths small-scale structures in these profiles. Both the depth of the velocity maximum, 1.1 MPa (110 m), and its association with the salinity maximum are typical near 0° N, 140° W. The stability frequency, N , is computed from the average temperature and salinity profiles and peaks slightly above the velocity maximum, producing a sharp maximum in the Richardson number, Ri .

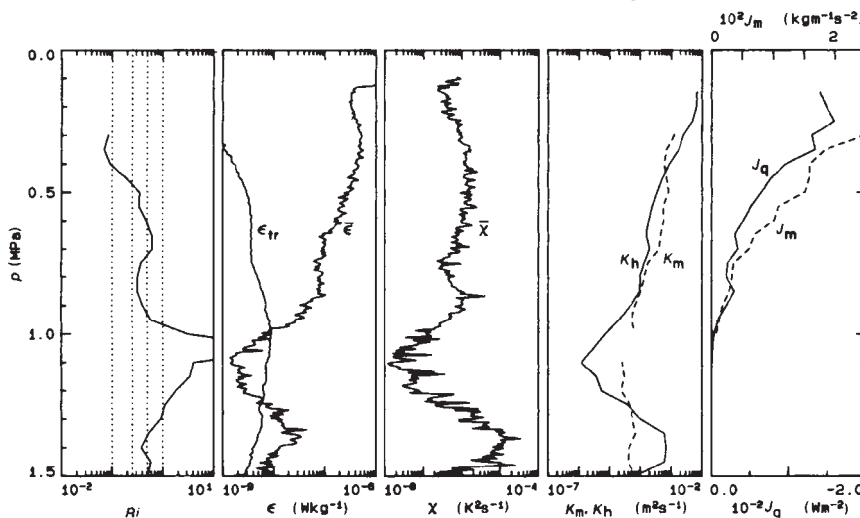
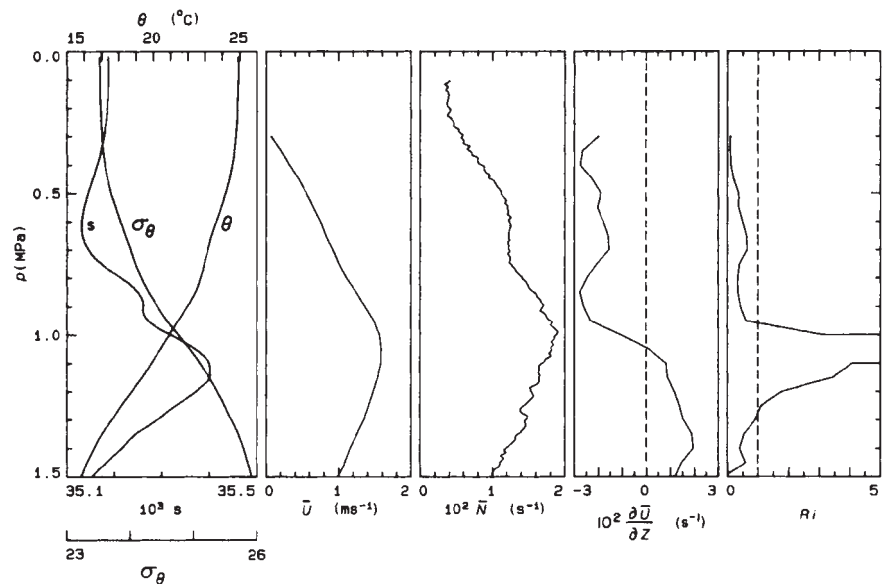


Fig. 2 The average Richardson number, Ri , and dissipation rates. Where $Ri < 1$, $\bar{\epsilon}(p) > \epsilon_{tr}$, the threshold for active turbulence. $\bar{\chi}(p)$ follows a similar pattern, modified by the local temperature gradient. The turbulent eddy coefficients for heat, K_h , and momentum, K_m (dashed line), are approximately equal between 0.3 and 0.9 MPa. The corresponding turbulent fluxes, J_q and J_m (dashed line), follow similar patterns, having large values near 0.3 MPa and decaying to about 0 at the Ri maximum.

same criterion holds when $Ri > 0.25$; when $Ri < 0.25$ the turbulence does not decay.

In the undercurrent, $\bar{\epsilon}$ decreases with depth (Fig. 2). Above 0.4 MPa, $Ri < 0.1$ and $\bar{\epsilon}$ exceeds ϵ_{tr} by at least two decades. This ratio decreases to about 20 between 0.6 and 0.9 MPa, where $1/4 < Ri < 1/2$. It then drops rapidly and crosses 1 where $Ri = 1$. Thus, in most of the high-shear zone above the velocity maximum the average dissipation rates indicate strongly active turbulence, producing a net buoyancy flux.

In steady, homogeneous turbulence, the balance equation for turbulent kinetic energy reduces to local production and dissipation. If the mean shear is strong, the turbulent production can be expressed as a vertical eddy coefficient for momentum, K_m

$$K_m = \frac{\bar{\epsilon}}{(1 - R_f) (\partial \bar{U} / \partial z)^2} \text{ m}^2 \text{ s}^{-1} \quad (1)$$

where R_f is the flux Richardson number. Because $R_f \leq 0.15$ this term is a small correction and is ignored. The corresponding vertical momentum flux is

$$J_m = -\rho K_m \frac{\partial \bar{U}}{\partial z} \text{ kg m}^{-1} \text{ s}^{-2} \quad (2)$$

In studies of the planetary boundary layer of the atmosphere, this procedure is known as the dissipation method and has been verified by comparison with direct flux measurements. If the temperature fluctuations are produced by turbulent overturns against the mean gradient, rather than by lateral advection, a similar balance applies with the $\bar{T}^2$ -balance equation, yielding

the turbulent transport coefficient for heat, K_h , and the vertical turbulent heat flux, J_q :

$$K_h = \frac{\chi}{2(\partial \bar{T} / \partial z)^2} \text{ m}^2 \text{ s}^{-1} \quad (3)$$

$$J_q = -\rho C_p K_h \frac{\partial \bar{T}}{\partial z} \text{ W kg}^{-1} \quad (4)$$

This approach, known as the Osborn-Cox method¹⁴, is widely used to obtain turbulent heat fluxes in the thermocline.

Equations (1) and (3) have been used with previous measurements in the undercurrent to develop parameterizations for numerical models^{15,16}. Their use has not been verified in the ocean by comparison with direct flux measurements, but conditions between 0.3 and 0.9 MPa are more appropriate to the dissipation method than any other known section in the ocean, and the present data offer much greater statistical confidence than previously available. Above 0.9 MPa, K_m and K_h are nearly equal, that is, the turbulent Prandtl number is close to one, and they decrease rapidly with depth. At 0.3 MPa, $K_m \approx K_h \approx 2 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$, and at 0.9 MPa, $K_m \approx K_h \approx 6 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$, a decrease by a factor of 25. Over the same interval, J_q drops from 140 to 20 W m^{-2} . (J_q decreases relatively less than K_h because $\partial \bar{T} / \partial z$ also decreases.)

Average mixing rates at velocity maximum

At the velocity maximum, the dissipation rate has a minimum, $\bar{\epsilon} = 2 \times 10^{-9} \text{ W kg}^{-1}$, which is about a quarter of ϵ_{tr} . Because $\bar{\epsilon} < \epsilon_{tr}$, dissipation-scale fluctuations are less likely to be isotropic

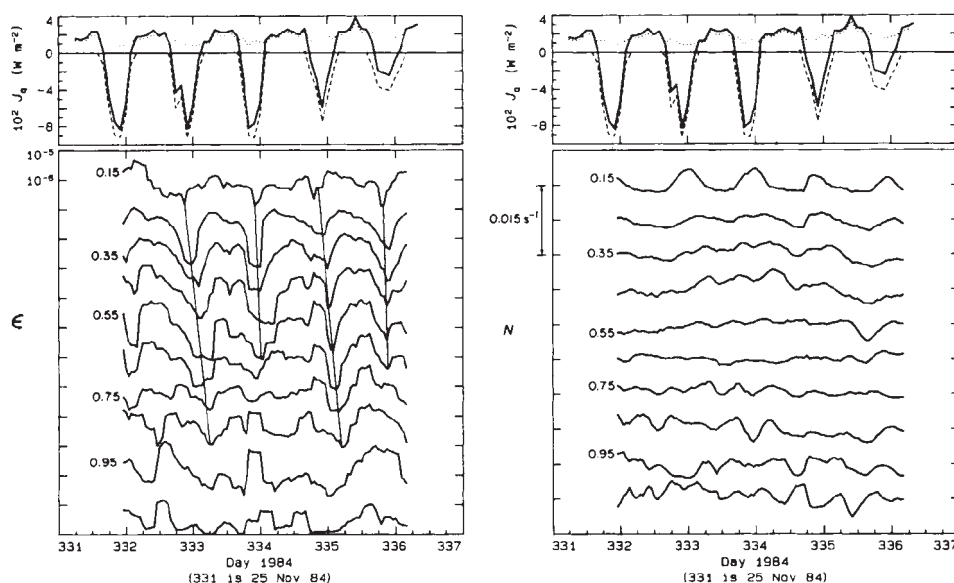


Fig. 3 Five-hour running means of the surface heat flux, of ϵ , and of N , showing a diurnal cycle in ϵ penetrating deep into the thermocline. The surface heat flux (solid line in the top panels) is dominated by the short-wave radiation (dashed), and by the latent heat flux (dotted). The time series for ϵ and N are averages over successive 0.1-MPa intervals; for example, the first set, labelled 0.15, are averages between 0.1 and 0.2 MPa. The curves for pressures >0.15 MPa are successively offset by one decade in ϵ and N , respectively. The diurnal cycle in ϵ remains distinct to 0.65 MPa and has some deeper signatures, unlike the diurnal N cycle, which is weak at 0.25 MPa and absent deeper. The daily minimum in ϵ has a progressive phase lag with depth, as shown by the slanting lines.

than at shallower depths. The maximum effect of anisotropy is to reduce $\bar{\chi}$ by a factor of 3. The effect on $\bar{\epsilon}$ is more difficult to estimate, but should be at least as large. Thus, if the microstructure at the core is anisotropic in the dissipation range, $\bar{\epsilon}$ and $\bar{\chi}$ will be even smaller compared with the levels in the high-shear zone shown in Fig. 2.

To remove the effects of displacements by internal waves and the internal tide, a preliminary statistical analysis was done on all 0.5-m ϵ values between potential density, $\sigma_\theta = 24.6$ and 25.5, which spans the velocity maximum. This yields $\bar{\epsilon} = 4.7 \times 10^{-9} \text{ W kg}^{-1}$ which is only a factor of three higher than an average found in a quiet region of the California Current¹⁷ and is almost three decades lower than the values reported from towed measurements^{6,7}. A probability distribution does not reveal any greater degree of intermittence than in other locations. For example, using the standard deviation of $\ln \epsilon$ as the measure of intermittence⁸, $\sigma_{\ln \epsilon} = 2.4$ compared with $\sigma_{\ln \epsilon} = 2.3$ from the California Current¹⁷. Using this in the maximum-likelihood estimate suggested for intermittent turbulence⁸ yields a value smaller than the arithmetic mean, that is, $\epsilon_{\text{mle}} = 4.2 \times 10^{-9} \text{ W kg}^{-1}$, because the distribution is skewed to low values and has a median of $2.2 \times 10^{-10} \text{ W kg}^{-1}$. Therefore, these data do not reveal unusual intermittence at the velocity maximum, but indicate that the conditions are similar to those in the main thermocline at less energetic locations.

Because the shear passes through zero, conditions are not appropriate for the dissipation method, and K_m is not calculated. The estimate of K_h is not affected, however, because the temperature gradient remains large, and $K_h \approx 1 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$. This is about 10 times κ_T , the molecular diffusion coefficient, and provides additional evidence that mixing levels at the velocity maximum are weak.

The diurnal cycle

Daily fluctuations, by factors of 10 to 100 in ϵ and χ , exist both in the diurnal mixed layer and in the well-stratified profile below (Fig. 3). The cycle extends to between 0.75 and 0.95 MPa, even though day-to-day variability not related to changes in the surface heat flux is also apparent.

Between 0.1 and 0.3 MPa, both the mixing and the stratification change in phase with the surface heat flux, which is dominated by the short-wave radiation and the latent heat flux. (The sensible heat flux is negligible by comparison, varying between -12 and $+20 \text{ W m}^{-2}$.) At night, the net flux is $\approx 200 \text{ W m}^{-2}$ out of the sea, corresponding to a surface buoyancy flux of $J_b^0 = 3.2 \times 10^{-7} \text{ W kg}^{-1}$, producing a well-mixed layer extending to between 0.2 and 0.3 MPa.

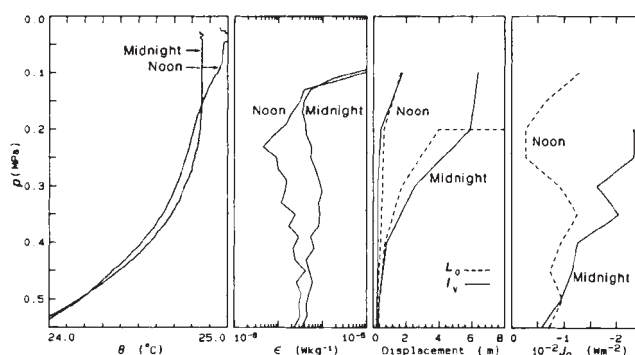
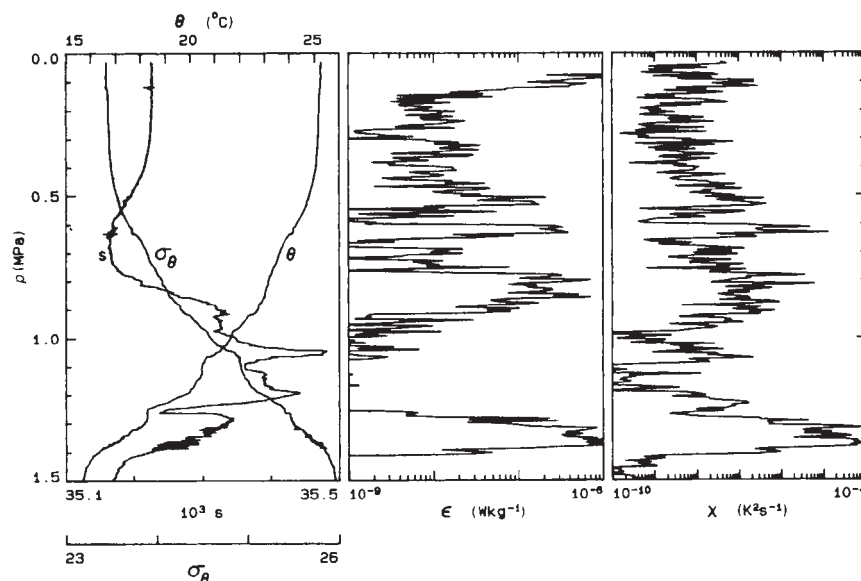


Fig. 4 Comparison of the averages of all profiles within 1 h of local noon and of local midnight. The mixed layer extends to 0.2 MPa at midnight, but is fully restratified by noon. Above 0.1 MPa the mixing rates are dominated by wind forcing, but between 0.15 and 0.65 MPa the diurnal cycle is evident. At noon, the turbulence is most heavily suppressed near 0.25 MPa. The observed vertical displacements, L_v , are independent measures of changes in ϵ and agree well with the Ozmidov scale, L_0 , computed from the dissipation rates and stratification. At 0.25 MPa, the turbulent heat flux, J_q , drops from 240 W m^{-2} at midnight to 30 W m^{-2} at noon.

Observations in another convecting mixed layer show $\bar{\epsilon} \approx J_b^0$ within the mixed layer and below the zone affected by the wind¹⁸. In this case, which was in a region with a moderate mean shear, the mixed layer terminated in a sharp density step and the dissipation rates below the entrainment zone decreased several decades. A profile of $\bar{\epsilon}$ for midnight shows a minimum of $4 \times 10^{-7} \text{ W kg}^{-1}$ below the surface zone, followed by an increase to $1 \times 10^{-6} \text{ W kg}^{-1}$ at 0.3 MPa (Fig. 4). Thus, the turbulence produced by convection may account for a substantial fraction of the observed ϵ above 0.3 MPa, but production by the mean shear or wind mixing is also present. The shape of the average midnight temperature profile is similar to individual realizations and shows a smooth transition to the stratified regime. This is very different from the abrupt termination of convecting mixed layers where the deeper water has low dissipation rates¹⁸. We believe this is due to the intense turbulence in the stratified section as well as in the mixed layer. Above 0.3 MPa, ϵ is a minimum at noon, when the profile has maximum stratification.

Below 0.3 MPa the diurnal ϵ cycle remains large and is 1.5 decades at 0.65 MPa (Fig. 3), even though there is day-to-day

Fig. 5 An AMP drop showing multiple salinity maxima and minima below 1 MPa, indicating the importance of lateral advection. The deepest of the three intrusions occurs where $Ri < 1$ and has strong mixing between 1.25 and 1.4 MPa, where the profile is diffusively unstable to salt fingering. The pattern of activity is similar to that found in highly sheared intrusions in a warm-core Gulf Stream ring¹⁶. This profile was taken at 20.53 GMT, or 10.53 LT, and has relatively weak dissipation rates between 0.3 and 0.9 MPa and on the intrusions where $Ri > 1$.



variability. The cycle is not symmetric, but has maxima lasting longer than the minima. Examination of more detailed plots shows that decreases in ϵ begin nearly simultaneously with depth. Nevertheless, the minimum has a phase lag that increases downward (Fig. 3). At 0.65 MPa, the lag is 3–6 h after local noon.

Between 0.3 and 0.9 MPa the diurnal cycles in ϵ and χ produce large changes in the overturning scales of the turbulence and in J_q (Fig. 4). Two measures of overturning are used: the Ozmidov scale, $L_o = \epsilon^{1/2} N^{-3/2}$, and l_r , the root-mean-square displacement computed by comparing the observed σ_θ profiles with resorted monotonic versions of the same data^{19,20}. The two scales show good agreement. Below 0.3 MPa, the overturning scales at noon are only a few tens of centimetres, but, during the night, they grow to a few metres. Correspondingly, J_q increases from about 30 W m^{-2} at noon to about 240 W m^{-2} at midnight. At night these estimates are unreliable above 0.25 MPa owing to the weak stratification in the mixed layer.

Thermohaline intrusions from 1.0 to 1.5 MPa

Between 1.0 and 1.5 MPa, thermohaline intrusions strongly affect ϵ and χ . Although temperature usually decreases monotonically over scales greater than 1 m, salinity has a sequence of maxima and minima (Fig. 5). These can only be of advective origin, and are probably associated with the salinity front near the Equator²¹. Changes between successive profiles indicate that the intrusions have horizontal scales less than a few kilometres. Very high dissipation rates occur beneath the salinity maximum of some intrusions. For example, comparing Fig. 5 with the average profiles in Fig. 2 shows that between 1.3 and 1.4 MPa ϵ is as high as $\bar{\epsilon}$ at 0.3 MPa, and χ is higher than $\bar{\chi}$. A similar pattern, found on thermohaline intrusions in a warm-core ring, was attributed to a combination of salt fingering and the shear of near-inertial motions²². Because strong lateral processes, and possibly salt fingering, are involved, neither the dissipation method nor the Osborn-Cox model are applicable to the mean profiles in this pressure range.

Discussion

Based on profiles of $\bar{\epsilon}$ and $\bar{\chi}$, five mixing zones are found in the upper 1.5 MPa below the near-surface zone: (1) the diurnal mixed layer, 0.1–0.3 MPa, has high dissipation rates, which are modulated by almost a factor of 100 and are in phase with the surface heat flux; (2) the upper high-shear zone, 0.3–0.9 MPa, has dissipation rates and turbulent fluxes that are high just below the mixed layer but decrease with depth, as Ri increases. At a given depth, the dissipation rates vary by factors of 10 to 100

in a diurnal cycle; (3) the velocity maximum, 0.9 to 1.25 MPa, has low dissipation rates, similar to those in mid-gyre; and (4) the lower high-shear zone, 1.25 to 1.5 MPa, has high dissipation rates when thermohaline intrusions occur.

The patterns and magnitudes of these observations are consistent with previous profiles in the undercurrent and are inconsistent with the towed observations. Discovery of the diurnal cycle in the stratified region below the surface mixed layer and of the role of thermohaline intrusions in the dissipation below the core is new.

The daily ϵ cycle is nearly as strong between 0.3 and 0.65 MPa as in the diurnal mixed layer. This is a surprise in view of the restricted depth range of the diurnal cycle in stratification. N variations are readily apparent (Fig. 3), but they have no diurnal signature similar to that in ϵ and χ . To our knowledge no previous observations have been reported, from any location, of a diurnal cycle below the mixed layer and its entrainment zone.

We have considered several possible causes of the deep diurnal mixing cycle: the absorption of solar radiation, a diurnal cycle in the large-scale shear, and daily modulation of high-frequency internal waves.

Suppression of the turbulence by a buoyancy flux resulting from the absorption of solar radiation seemed a likely mechanism when we noticed the diurnal cycle at sea and has since been suggested to us by several others (M. Lewis and J. D. Woods, personal communications). We do not, however, understand how this can happen without also producing a daily cycle in N , which is not apparent in Fig. 3 or in other more detailed plots.

No diurnal shear cycle has been reported from the Equator; but, in response to discovery of the diurnal mixing cycle, current meter records from moorings at 0°N , 140°W in May and November 1983 have been examined (D. Halpern, personal communication). Between 0.1 and 0.25 MPa, shear has a strong daily fluctuation, $\sim 0.01 \text{ s}^{-1}$, but this decreases rapidly with depth and is negligible below the diurnal mixed layer, especially compared with the semi-diurnal tidal signal.

Internal wave energy on the equator is higher than found in mid-gyre²³, and laboratory studies show that mixed layer turbulence can generate internal waves in the thermocline²⁴. Therefore, internal waves, generated when turbulence is strong in the mixed layer, may propagate downward and break in the stratified high-shear zone, producing much of the observed mixing. When the surface layer restratifies during the day the source of internal waves will be suppressed, leading to a minimum in mixing below the mixed layer which may have a phase lag as the effect propagates downward.

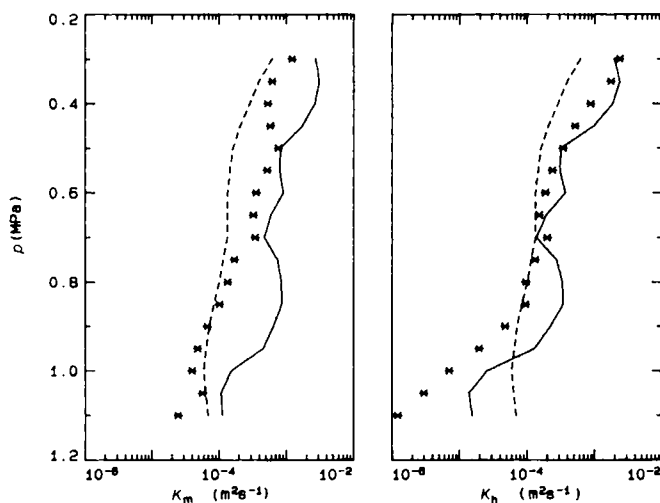


Fig. 6 Comparison of K_m and K_h obtained using the dissipation method (*) with parameterizations used in two numerical models of the undercurrent. Dashed line, from ref. 15; solid line, from ref. 16. The model K_m and K_h are functions of the observed $\bar{N}(p)$ and/or $Ri(p)$. The parameterizations are closer where Ri is low, above 0.7 MPa, than where Ri is high.

The data available to us are inadequate to determine whether one of these mechanisms is the cause, because small changes in either N or shear could be responsible. Mixing in this zone resembles a sharply tuned harmonic oscillator, which can have large output changes for small forcing perturbations. Plotting the $\bar{\epsilon}$ profile as a function of Ri shows a large increase when Ri drops below 0.25. Consequently, a 20% increase in shear or a similar decrease in N can change the average Richardson number from 0.35 to 0.25. Figure 3 shows much larger changes occurring in N , but with no apparent daily frequency. The eight XCP profiles show that the shear also changes by more than 20%. Because a wide band of frequencies seem to contribute to changes in both N and shear, longer time series of both are needed to define a diurnal signature. Data taken by other Tropic Heat investigators may be able to resolve this, but we know of no measurements that can examine the role of high-frequency internal waves and recommend that such measurements be included in future experiments.

In addition to taking the time series on the Equator we made measurements at 0.5° intervals off the Equator and found that the diurnal cycle extended to at least 1.0 MPa just outside of the undercurrent, compared with 0.3 MPa, in the centre of the undercurrent. It is possible that internal waves radiated from these deeper layers are also involved and that the problem is three-dimensional rather than just one-dimensional.

Two recent numerical models of the undercurrent are based on previous microstructure observations. One parameterizes the eddy coefficients in terms of N , largely for mathematical convenience¹⁵

$$K_m = K_h = AN^{-2} \quad (5)$$

with A chosen to produce $K_m = 5.5 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ at N_{max} . Another¹⁶ bases the parameterization on Ri

$$K_m = 1 \times 10^{-4} + \frac{5.5 \times 10^{-3}}{(1 + 5 Ri)^2} \quad (6)$$

and

$$K_h = 1 \times 10^{-5} + \frac{K_m}{1 + 5 Ri} \quad (7)$$

Comparison with our estimates, using equations (1) and (3), shows that equation (6) is consistently high, except between 0.5 and 0.7 MPa (Fig. 6). Equation (7) agrees remarkably well with observations above 0.7 MPa, but is an overestimate at greater depths. Equation (5) agrees relatively well with observations for K_m , but is low by factors of 2–5 between 0.4 and 0.7 MPa. For K_h , equation (5) has much less vertical change than the observations, and crosses from an under to an over estimate near 0.7 MPa. At 1.1 MPa, it is almost a factor of 100 too high.

Both parameterizations are in rough agreement with K_m and K_h computed from our data and equations (1) and (3), but neither is a good fit. The major discrepancies are overestimation of both eddy coefficients at the velocity maximum (the additive terms are much too large) and underestimation of the increases in K_m and K_h when Ri drops below 0.25. A plot of K_m and K_h as functions of Ri shows that they rise abruptly by about a factor of 10 when Ri approaches 0.25. Because a large section of the mixing region is in this Richardson number range, improving this part of the parameterization is particularly important.

Until the diurnal cycle is understood, these parameterizations, which are based on only the large-scale profiles, must be viewed as tentative. If fluctuating internal wave energy or the absorption of solar energy has a major role, a simple Ri parameterization will be inadequate. Discovery of the diurnal cycle also raises the possibility of an annual cycle, because the intensity of the diurnal cycle may vary with season.

We thank Captain Clappitt and the officers and crew of the RV *Thomas G. Thompson* for their cooperation at sea and Wayne Nodland, Dale Hirt, Earl Krause, Matt Nicholas, Pat McKeown, Steve Sova and Thomas Bongers for careful work in collecting the data, and also David Halpern for examining his previous equatorial shear data. This work was funded by the NSF on grant OCE-8214780, as part of the Tropic Heat Program. School of Oceanography, University of Washington, contribution no. 1597.

Received 5 June; accepted 28 August 1985.

- Eriksen, C. C. *EOS* **66**, 50–51 (1985).
- Gregg, M. C. *J. geophys. Res.* **81**, 1180–1196 (1976).
- Crawford, W. R. & Osborn, T. R. *Deep-Sea Res. GATE suppl.* 2 to vol. **26**, 285–308 (1979).
- Crawford, W. R. *J. phys. Oceanogr.* **12**, 1137–1149 (1982).
- Paka, V. T., Vasilenko, V. M., Osadchii, A. S. & Shkurenko, V. I. in *Variability of the Ocean and Atmosphere in the Equatorial Atlantic* (ed Monin, A. S.) 78–105 (Nauka, Moscow, 1982).
- Belyayev, V. S., Lyubimtsev, M. M. & Ozmidov, R. V. *Izv. Atmos. Ocean. Phys.* **9**, 1179–1185 (1973).
- Williams, R. B. & Gibson, C. H. *J. phys. Oceanogr.* **4**, 104–108 (1974).
- Gibson, C. H. in *Hydrodynamics of the Equatorial Ocean* (ed. Nihoul, J. C. J.) 131–153 (Elsevier, Amsterdam, 1983).
- Gregg, M. C., Nodland, W. E., Aagaard, E. E. & Hirt, D. H. *Oceans '82 Conf. Rec. IEEE Cat. No. 82CH1827-5*, 260–265 (Marine Technology Society, Washington DC 1982).

- Sanford, T. B., Drever, R. G., Dunlap, J. H. & D'Asaro, E. A. *Design, Operation and Performance of an Expendable Temperature and Velocity Profiler*, APL-UW 8110 (Applied Physics Laboratory, University of Washington, Seattle, 1982).
- Howard, L. N. *J. Fluid Mech.* **10**, 509–512 (1961).
- Stillinger, D. C., Helland, K. N. & Van Atta, C. W. *J. Fluid Mech.* **131**, 91–122 (1983).
- Rohr, J. J., Helland, K. N., Itsweire, E. C. & Van Atta, C. W. *5th Symp. on Turbulent Shear Flows*, Cornell (submitted).
- Osborn, T. R. & Cox, C. S. *Geophys. Fluid Dyn.* **3**, 321–345 (1972).
- McCreary, J. P. *Phil. Trans. R. Soc. A298*, 603–635 (1981).
- Pacanowski, R. C. & Philander, S. G. H. *J. phys. Oceanogr.* **11**, 1443–1451 (1981).
- Gregg, M. C., D'Asaro, E. A., Shay, T. J. & Larson, N. J. *J. phys. Oceanogr.* (submitted).
- Shay, T. J. & Gregg, M. C. *Nature* **310**, 282–285 (1984).
- Thorpe, S. A. *Phil. Trans. R. Soc.* **286**, 125–181 (1977).
- Dillon, T. M. *J. geophys. Res.* **87**, 9601–9613 (1982).
- McPhaden, M. J. *J. geophys. Res.* (in the press).
- Larson, N. G. & Gregg, M. C. *Nature* **306**, 26–32 (1983).
- Hayes, S. P. & Powell, C. J. *J. geophys. Res.* **85**, 4029–4035 (1980).
- Linden, P. F. *J. Fluid Mech.* **71**, 385–405 (1975).

Hotspot magmas can form by fractionation and contamination of mid-ocean ridge basalts

Don. L. Anderson

Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Melts ascending from deep reservoirs and trapped beneath thick lithosphere exchange heat with and partially melt the shallow mantle as they crystallize. Basalts erupting through thick lithosphere are more likely to have experienced deep crystal fractionation and contamination than those at mature spreading centres. Geochemical data for ocean island and other hotspot magmas do not require a primitive or lower mantle source.

MID-OCEAN ridge basalts (MORB) are the most depleted (low concentrations of large-ion lithophile (LIL) elements, low values of Rb/Sr, Nd/Sm, $^{87}\text{Sr}/^{86}\text{Sr}$, $^{144}\text{Nd}/^{143}\text{Nd}$, $^{206}\text{Pb}/^{204}\text{Pb}$) and the most voluminous magma type. They erupt through thin lithosphere and have apparently experienced some crystal fractionation before eruption. Isotope ratios show that the MORB reservoir was outgassed and depleted more than 10^9 yr ago. There is growing evidence that there is also an ancient enriched reservoir in the mantle and that the continental crust is not the only complement to the MORB reservoir¹⁻⁴. The most enriched mantle magmas are lamproites, kimberlites, and basalts from islands such as Kerguelen, St Helena and Tristan da Cunha. There is a complete spectrum of basalts lying between these extremes. The presence of ^3He in MORB and the fact that its $^3\text{He}/^4\text{He}$ is higher than atmospheric^{5,6} suggests that outgassing was not 100% efficient, or that MORB is contaminated. Pb isotopes suggest that MORB has experienced some contamination before eruption⁷. Therefore, MORB itself may be a hybrid magma.

The geochemistry of various magma types requires at least two distinct reservoirs in the mantle. The relatively uniform depleted mantle (DM) suffered an ancient depletion event, presumably by removal of a residual fluid. The other reservoir, which we call enriched mantle (EM), is less homogeneous and is presumed to provide the enriched signature of hotspot magmas. Correlated LIL and isotope variations in mantle magmas are often explained by mixing of material from the EM and DM reservoirs.

Basalts with high Rb/Sr, La/Yb and Nd/Sm ratios generally have high $^{87}\text{Sr}/^{86}\text{Sr}$ and low $^{143}\text{Nd}/^{144}\text{Nd}$ ratios. These can be explained by binary mixing of depleted and enriched magmas or by mixtures of a depleted magma and a component representing varying degrees of melting of an enriched reservoir⁷. The variable LIL ratios of such an enriched component generate a range of mixing hyperboles or 'scatter' about a binary mixing curve, even if there are only two isotopically distinct end members. Thus a model with two isotopically distinct reservoirs can generate an infinite variety of mixing lines⁷. In some regions, however, the inverse relationship between LIL and isotope ratios cannot be explained by binary mixing⁸⁻¹³. These regions are all in midplate or thick lithosphere environments and sublithospheric crystal fractionation involving garnet and clinopyroxene might be expected before eruption.

The wide range of isotope ratios in island and continental basalts has led to the view that the mantle is heterogeneous on all scales. The failure of simple binary mixing to explain the spread of isotope and trace element ratios, and the presence of LIL-enriched basalts with time-integrated depleted isotope ratios, has also given rise to the concept that recent mantle metasomatism has affected the source region and may even be a prerequisite for magmatism¹⁴.

Some of the observations which have been interpreted in terms of metasomatism, or an extremely heterogeneous mantle are: (1) scatter of $^{143}\text{Nd}/^{144}\text{Nd}$ - $^{87}\text{Sr}/^{86}\text{Sr}$ values about the mantle array; (2) lack of correlation between $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{206}\text{Pb}/^{204}\text{Pb}$

or $^4\text{He}/^3\text{He}$; and (3) inverse correlations between La/Ce or Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$. However, if one or both end members of a hybrid is a fractionated melt, or represents varying degrees of partial melting, then these observations do not necessarily require multiple sources. If melts from one reservoir must traverse the other reservoir on their way to the surface there is no need for a 'plum pudding' mantle or recent metasomatism of the source.

The cases where a depleted magma causes variable degrees of melting of an enriched reservoir or an undepleted magma partially melts a depleted layer have already been treated^{7,8}. We now treat a fractionating depleted magma interacting with an enriched reservoir. As a point of reference, we consider the depleted magma from DM to be the parent of MORB. If this magma is brought to a near-surface environment it may fractionate olivine, plagioclase and orthopyroxene. If arrested by thick lithosphere it will fractionate garnet (gt) and clinopyroxene (cpx). We assume that the fractionating crystals and melt are in equilibrium. We emphasize varying degrees of crystal fractionation but our results are applicable to the case where the depleted component represents varying degrees of partial melting⁸ of a gt- and cpx-rich region, such as garnet pyroxenite lower oceanic lithosphere, or an eclogite slab.

There are two situations that have particular relevance to mantle-derived magmas. Consider a 'normal' depleted mantle magma. If it can rise unimpeded from its source to the surface, such as at a rapidly spreading ridge, we have a relatively unfractionated, uncontaminated melt, that is, MORB. Suppose now that the magma rises in a midplate environment, and its ascent is impeded by thick lithosphere. The magma will cool and crystallize, simultaneously partially melting or reacting with the surrounding shallow mantle. Thus, crystal fractionation and mixing occur together, and the composition of the hybrid melt changes with time and with the extent of fractionation. Fractionation of gt and cpx from a tholeiitic or picritic magma at sublithospheric depths (>50 km) can generate alkalic magmas with enriched and fractionated LIL patterns. We investigate the effects of combined eclogite fractionation (equal parts of gt and

Table 1 Parameters of endmembers

	Depleted mantle 1	Enriched mantle 2	Enrichment ratio	
$^{87}\text{Sr}/^{86}\text{Sr}$	0.701, 0.702	0.722	Sr_2/Sr_1	13.8
$^{143}\text{Nd}/^{144}\text{Nd}$	24.6	-16.4	Nd_2/Nd_1	73.8
$^{206}\text{Pb}/^{204}\text{Pb}$	16.5, 17.0	26.5	Pb_2/Pb_1	62.5
$^3\text{He}/^4\text{He}$	6.5 R_a	150 R_a	He_2/He_1	2
La/Ce	0.265	0.50	Ce_2/Ce_1	1 65
Sm/Nd	0.50, 0.375	0.09, 0.139		
Partition coefficients				
Rb	0.02	Sr 0.04	La	0.012
Ce	0.03	Nd 0.09	Sm	0.18
Pb	0, 0.002	He 0.50		

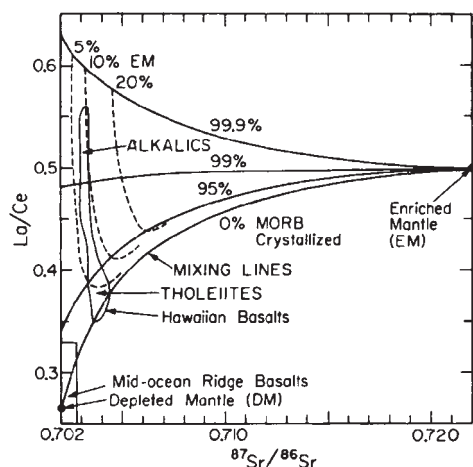


Fig. 1 La/Ce versus $^{87}\text{Sr}/^{86}\text{Sr}$ for the fractionation-contamination model, compared with Hawaiian basalts. Unfractionated MORB has La/Ce = 0.265. La/Ce of the depleted endmember increases as crystal fractionation proceeds. The enriched endmember has La/Ce = 0.5, in the range of kimberlitic magmas. The Hawaiian tholeiites can be modelled as mixes ranging from pure MORB plus 2–7% enriched component to melts representing residuals after 95% cpx plus gt crystal fractionation and 5–8% enriched component. Alkali basalts involve more crystal fractionation and more contamination. In all the figures, solid curves are mixing lines between EM and melts representing fractionating depleted magmas. Dashed curves are trajectories of constant mixing proportions. Data from ref. 8.

cpx) and 'contamination' on melts from DM. 'Contamination' is modelled by mixing an enriched component with the fractionating depleted magma. This component is viewed as a partial melt generated by the latent heat associated with the crystal fractionation. The assumed geochemical properties of the endmembers, the enrichment factors of the elements in question and the partition coefficients, D , assumed in the modelling are given in Table 1. D for Rb, Sr, Sm, Nd, La and Ce are the mean for cpx and gt¹⁵. Later we shall discuss Pb and He and, in these cases, D is an adjustable parameter. The figures show various ratios for mixes of a fractionating depleted melt and an enriched component (EM). I assume equilibrium crystal fractionation as appropriate for a turbulent or permeable magma body, and constant D . I will not attempt to fit the data displayed in the figures. My purpose is simply to show the range of parameters required to span the data by the proposed mechanism. I do not rule out other contributions to the geochemical complexity of magmas. In particular, I shall discuss the heterogeneity of EM.

La/C versus $^{87}\text{Sr}/^{86}\text{Sr}$

La/Ce is high in melts relative to crystalline residues containing gt and cpx. It is low in depleted reservoirs and high in enriched reservoirs. Low and high values of $^{87}\text{Sr}/^{86}\text{Sr}$ are characteristics of time-integrated depleted and enriched magmas, respectively. As high $^{87}\text{Sr}/^{86}\text{Sr}$ implies a time-integrated enrichment of Rb/Sr, there is generally a positive correlation of Rb/Sr and La/Ce with $^{87}\text{Sr}/^{86}\text{Sr}$. Some magmas⁹, however, exhibit high La/Ce and low $^{87}\text{Sr}/^{86}\text{Sr}$. This cannot be explained by binary mixing of two homogeneous magmas.

On a theoretical La/Ce versus $^{87}\text{Sr}/^{86}\text{Sr}$ plot (Fig. 1) the mixing lines between the crystallizing MORB and EM components reverse slope when MORB has experienced slightly more than 99% crystal fractionation. The relationships for equilibrium partial melting and equilibrium, or batch, crystallization are, of course, the same. Therefore, large degrees of crystallization of a MORB-like melt or small amounts of partial melting of a MORB layer are implied by an inverse relationship between La/Ce (or Rb/Sr, La/Yb, Nd/Sm, and so on) and $^{87}\text{Sr}/^{86}\text{Sr}$ (or $^{144}\text{Nd}/^{143}\text{Nd}$) such as observed at many mid-plate

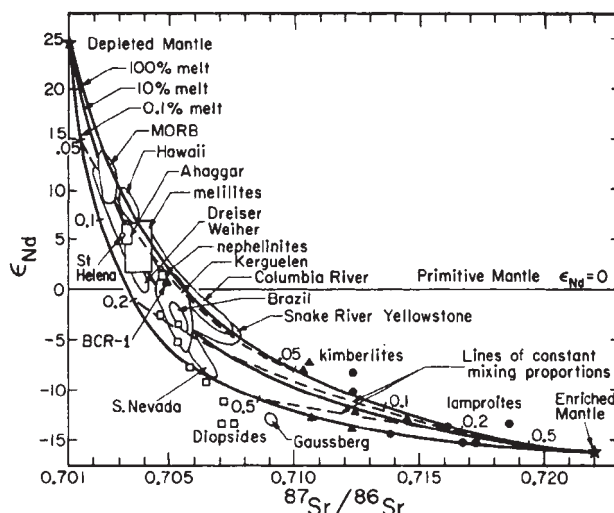


Fig. 2 ϵ_{Nd} versus $^{87}\text{Sr}/^{86}\text{Sr}$ for mixtures involving a depleted magma or residual fluids from such a magma after crystal fractionation, and an enriched component (EM). Data from various sources including refs 4, 38–44.

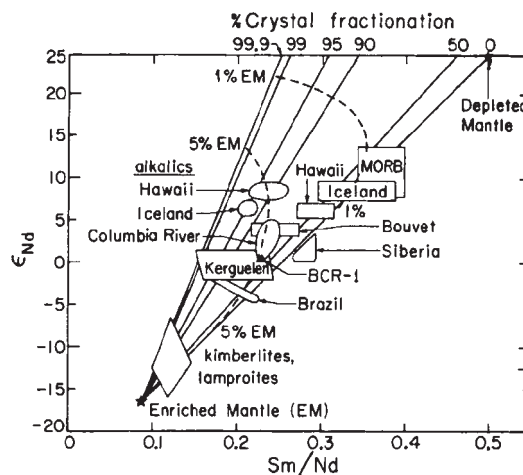


Fig. 3 $^{143}\text{Nd}/^{144}\text{Nd}$ (relative to primitive mantle) versus Sm/Nd for mixtures of enriched mantle and residual melts resulting from high-pressure crystal fractionation of MORB-like depleted magma. Data from refs 8, 15, 38–40. See ref. 45 for supporting evidence for Iceland.

environments. The apparently contradictory behaviour of magmas with evidence for current enrichment and long-term depletion is often used as evidence for 'recent mantle metasomatism'. Figure 1 illustrates an alternative explanation. Note that, with the parameters chosen, the Hawaiian alkalis have up to 10% contamination by the enriched component.

ϵ_{Nd} versus $^{87}\text{Sr}/^{86}\text{Sr}$

Isotope ratios for ocean island and continental basalts are compared with mixing curves in Fig. 2. These basalts can be interpreted as mixes between a fractionating depleted magma and an enriched component. The value for primitive mantle ($\epsilon_{\text{Nd}} = 0$) is also shown. The primitive mantle value of $^{87}\text{Sr}/^{86}\text{Sr}$ is unknown and cannot be inferred from basalts that are themselves mixtures. For example, the basalts cover a field extending from 0 to 99.9% crystal fractionation and the corresponding $^{87}\text{Sr}/^{86}\text{Sr}$ ratios at $\epsilon_{\text{Nd}} = 0$ range from 0.7043 to 0.7066. The extreme situations, constant degree of fractionation combined with variable contamination, and constant mixing proportions combined with variable fractionation, can each explain the trend of the data. Together, they can explain most of the dispersion of the data.

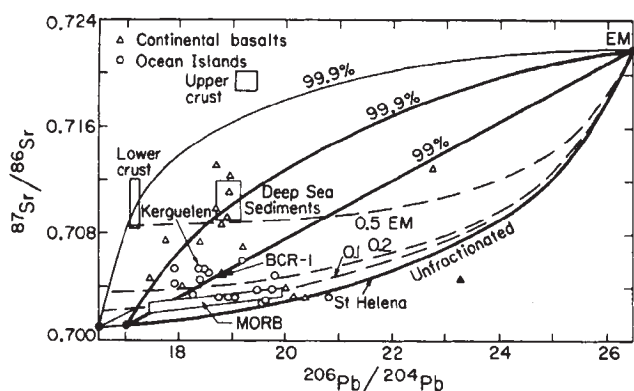


Fig. 4 $^{87}\text{Sr}/^{86}\text{Sr}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ trajectories for mixtures of a fractionating depleted MORB-like magma and an enriched component. Mixing hyperboles are solid lines. Hawaiian basalts fall in the region of 50–99% crystal fractionation (garnet plus clinopyroxene) and 10–30% contamination (EM). Data from various sources including refs 23, 46–49.

ϵ_{Nd} versus Sm/Nd

The mixture-fractionation curves for this system are shown in Fig. 3. High Sm/Nd basalts from Iceland, Hawaii, Siberia, Kerguelen and Brazil all fall near the curve for unfractionated MORB with 1–5% contamination. Alkalics from large oceanic islands with thick crust (Hawaii, Iceland and Kerguelen) are consistent with large amounts of crystal fractionation and moderate (5–10%) amounts of contamination.

The interpretation is that the more voluminous tholeiites comprise slightly fractionated and contaminated MORB, while the alkalics have experienced sublithospheric crystal fractionation and contamination before eruption. MORB itself has ~1% contamination, similar to that required to explain Pb isotopes⁷. The Hawaiian tholeiites can be modelled as MORB that has experienced variable degrees of deep crystal fractionation (0–95%), mixed with 1–5% of an enriched component. The alkali basalts represent greater extents of crystal fractionation and contamination. The Columbia River basalts can be interpreted as MORB which has experienced 80% crystal fractionation and 5% contamination.

Lead and helium isotopes

Rb, Sr and the light rare earth elements (REE) are classic incompatible elements and the effects of partial melting, fractionation and mixing can be explored with some confidence. Relations between these elements and their isotopes should be fairly coherent. Some enriched magmas also have high $^3\text{He}/^4\text{He}$, $\delta^{18}\text{O}$ and $^{206}\text{Pb}/^{204}\text{Pb}$. These isotopes provide important constraints on mantle evolution, but they may be decoupled from the LIL variations and are usually assumed to be so. $^3\text{He}/^4\text{He}$ depends on the outgassing and magmatic history of the depleted reservoir and the U and Th content and age of the enriched component. $^{206}\text{Pb}/^{204}\text{Pb}$ and $^3\text{He}/^4\text{He}$ are sensitive to the age of various events affecting a reservoir because of the short half life of U. The U/Pb ratio may be controlled by sulphides and metals as well as silicates. Nevertheless, it is instructive to investigate the effects of contamination and fractionation on these systems to seek out what other effects or components are involved.

There is no simple relationship between the Sr and Pb isotopes in mantle magmas. A minimum of three components is required to satisfy both isotope systems when only simple mixing is considered. St Helena, for example, has high $^{206}\text{Pb}/^{204}\text{Pb}$ –low $^{87}\text{Sr}/^{86}\text{Sr}$, Kerguelen has high $^{87}\text{Sr}/^{86}\text{Sr}$ –low $^{206}\text{Pb}/^{204}\text{Pb}$, and MORB has low $^{87}\text{Sr}/^{86}\text{Sr}$ –low $^{206}\text{Pb}/^{204}\text{Pb}$. The Pb isotopes in MORB are higher than expected for a primitive or depleted

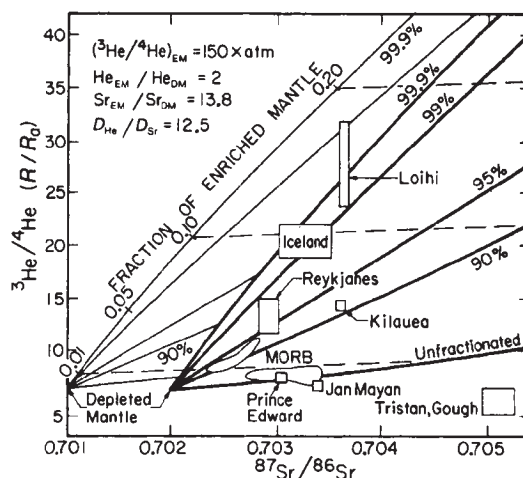


Fig. 5 Mixing curves for $^3\text{He}/^4\text{He}$ versus $^{87}\text{Sr}/^{86}\text{Sr}$. Two choices for the depleted endmember are shown. Hawaiian basalts imply 90 to >99% crystal fractionation and 5–18% contamination. To make Iceland consistent with Fig. 3, one could raise Sm/Nd (DM) or R_a (EM). Data from refs 50–53.

reservoir. Figure 4 shows a mixing-fractionation curve for $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{206}\text{Pb}/^{204}\text{Pb}$. To span the data, with only two distinct endmembers, it has to be assumed that the partition coefficient for Pb is much less than that for Sr. This would be the case if Pb is entirely contained in low-melting-point phases or is much more incompatible in gt and cpx than is Sr. Figure 4 shows that the data can be explained with only two isotopically distinct mantle reservoirs. However, even if EM is global it can have variable $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ as these are sensitive to the U/Pb ratio and the age of enrichment events. U/Pb and $^3\text{He}/\text{U}$ of EM will change locally by the subduction of sediments and oceanic crust and by the insertion of basalts which do not make their way to the surface. Isotope ratios of Pb and He, therefore, are variable and change rapidly with time. The various contributions to the enrichment process (subduction, trapped magmas) make it likely that EM, the shallow mantle in our case, is laterally inhomogeneous.

Mid-ocean ridge tholeiites¹⁶ average about 8.5 times the atmospheric ratio (R_a) for $^3\text{He}/^4\text{He}$. Samples from Hawaii^{17,18} have R_a up to 37 and natural diamonds^{19,20} range up to 280. $^3\text{He}/^4\text{He}$ near Iceland and Hawaii can be modelled as two-component mixtures with MORB and Loihi being the endmembers. Other islands have lower ratios than MORB, suggesting a third component, a variably outgassed EM or reduction in $^3\text{He}/^4\text{He}$ by U–Th decay. The high $^3\text{He}/^4\text{He}$ of some ocean island basalts is often assumed to require decoupling from other processes and a lower mantle origin.

Results are shown in Fig. 5. The highest $^3\text{He}/^4\text{He}$ samples from Loihi imply ~18% of the enriched endmember ($R_a = 150$) and >99% crystal fractionation. Prince Edward and Jan Mayen can be modelled as mixtures involving nearly unfractionated MORB. Basalts from Tristan and Gough may require a third component, such as subducted sediments, or they may have evolved in a high U–Th environment. To cover the total range with the chosen endmembers, it is necessary to assume that the He abundance in EM is about twice the abundance in the depleted magma and that D for He is much greater than that for Sr. The partition coefficient for He is unknown, but it may not behave as a classic incompatible element. Clinopyroxene–melt ‘partition’ (or ‘trapping’) coefficients for Ar and Kr are 0.2–0.5, and may increase with decreasing atomic weight²⁰. These surprisingly high values are of the order required to explain the He–Sr isotope results. The distribution of $^3\text{He}/^4\text{He}$ in the enriched reservoir is likely to be uneven because it depends on the amount of trapped ^3He , the U and Th content, and the age of enrichment. If part of the LIL enrichment is due to reinjection of oceanic crust and sediments, this part of EM will have a low

$^3\text{He}/^4\text{He}$ ratio. Ancient enriched reservoirs will have higher $^{206}\text{Pb}/^{204}\text{Pb}$ and lower $^3\text{He}/^4\text{He}$ ratios than equivalent modern reservoirs. Nevertheless, combined crystal fractionation and contamination, using only two endmembers, can explain much of the range and several of the trends observed in oceanic basalts.

Discussion

Chen and Frey⁸ propose a model for Hawaiian volcanism that involves melting of an enriched mantle plume from the lower ('primitive') mantle. This upwelling plume traverses a depleted (MORB) region, which partially melts (0.05–2.0%); these melts mix in varying proportions to form the spectrum of basalts found in Hawaiian volcanoes. With their choice of endmembers the MORB component must be small, <2%.

Our results suggest an alternative explanation. Melts from DM rise to the base of the lithosphere, fractionate and mix with melts from EM. At mid-ocean ridges the lithosphere is thin, crystal fractionation occurs at shallow depths, and eruption occurs with little contamination. Mid-plate volcanism involves crystal fractionation at greater depth, >60 km. Because of slower cooling at depth, there is more chance for mixing with melts from the asthenosphere and lithosphere before eruption. Thus, diapirs from the depleted reservoir may be the precursors for both mid-ocean and mid-plate volcanism, but the trace-element and isotope spectra of basalts depend on the thickness of the lithosphere or crust. Material from DM, which initiates melting in the shallow enriched reservoir, dominates the mix. The choice of a single EM component is simplistic but it illustrates the mechanism. On their ascent, MORB diapirs can interact with asthenosphere, oceanic or continental lithosphere, sediments, sea water and subducted material.

In the Chen–Frey mode, the amount of the MORB component decreases as the inferred degree of partial melting of the MORB source increases, the reverse of what might be expected. The relative portions of the MORB and enriched components are also the reverse of what the $^3\text{He}/^4\text{He}$ data indicate. Kaneoka²¹ suggests that the MORB component of Hawaiian basalts represents more than 50–70%. In the present model, the amount of contamination increases as crystal fractionation proceeds.

Kaneoka²¹ favours a model in which LIL- and ^3He -rich material from a deep primitive mantle reservoir rises through and mixes with a shallow, depleted region. He argues that if the enriched layer is shallow, plumes should originate at shallow depth, and the only overlying depleted layer is the oceanic lithosphere. However, the distinctive geochemical signature of hotspot basalts may be acquired at shallow depths, and most of the erupted material may be depleted basalts rising from greater depth. Continental contamination is an oft-invoked mechanism of this type, although it does not work for the rare gases. An LIL- and ^3He -rich shallow mantle does not imply that plumes originate at shallow depth. In discussions of hotspots, it should be recognized that the source of the heat, the source of most of the magma and the source of the isotope signatures may be at different depths. Hot regions of the lower mantle, for example, can localize melting in the transition region (400–650 km depth) and depleted melts from this region may become contaminated as they rise. Melts trapped in, or rising through, the shallow mantle may be responsible for the distinctive physical properties of the low-velocity zone and asthenosphere. Most workers assume that noble gases are strongly partitioned into melts which then efficiently lose these gases to the atmosphere. Outgassing, however, probably occurs only at relatively shallow depths, where the major gases, such as CO_2 and H_2O , can exolve²². If melts are trapped at depth, the noble gases will be as well. The presence of ^3He and ^{36}Ar in magmas does not require the existence of a previously unmelted or primitive reservoir. The $^3\text{He}/^4\text{He}$ characteristics of hotspot magmas, therefore, do not require an explanation separate from that for the other isotopes, except that $^3\text{He}/^4\text{He}$ can decrease rapidly with time in a high-U reservoir.

The magnitude of tholeiitic volcanism, the dominance of the MORB component in mantle magmas and the close temporal

and spatial relationship of depleted and more enriched and evolved alkalic magmas support a controlling role for melts from DM in mantle magma genesis. I propose that oceanic ridges, islands and seamounts share a common source and that mantle magmas differ primarily because of different degrees of sublithospheric crystal fractionation and contamination, and the nature of the (minor) enriched component. Shallow fractionation, crustal contamination, alteration and recycling, will be superimposed on the effects considered here. There is also the possibility that EM is inhomogeneously enriched or veined. Pb isotopes, for example, require more than two endmembers²³. These may differ in age or intrinsic chemistry, or both.

The location of EM, the source of the contamination discussed here, cannot be settled by petrological and geochemical arguments alone. It has been a common, but unfortunate, practice in the recent geochemical literature to avoid the issue by using the terms lower mantle, primitive mantle, undegassed mantle, plume reservoir and ocean island reservoir, interchangeably. The similarity in trace-element and isotope geochemistry between large islands, seamounts, continental rifts and island arcs suggests that EM is widespread and shallow. The most enriched magmas, kimberlites, appear to originate at a depth of ~200 km. Mid-ocean ridges can be traced by seismic means to depths >200 km, in some cases, >400 km (refs 24, 25). The ascent phase of a deep diapir is characterized by melting. Crystallization is caused by entry of the diapir into the cold surface boundary layer. The heat of crystallization causes melting, or wall-rock reactions, and, hence, crystal fractionation and contamination are concurrent processes. The adoption of a single-stage process involving two isotopically distinct endmembers, only one of which is allowed to fractionate, is clearly an oversimplification. The enriched component may represent variable degrees of partial melting of an enriched region which is probably inhomogeneous. The shallow mantle is periodically replenished with subducted material and trapped magmas. Some of these processes serve to diminish the extreme amount of crystal fractionation implied by the present calculations, as can non-Henry's Law behaviour^{26,27}. The possibility of large amounts of eclogite or clinopyroxene fractionation has been discussed by O'Hara²⁸ and others^{29–36}.

There are several mechanisms by which EM might be formed: (1) Melts removed from DM are distributed between the crust and the shallow mantle. (2) Rejection of sediments and crust (recycling). In both cases, EM is expected to be shallow. The melt and crustal extraction processes are not 100% efficient; thus, LIL-rich melts are trapped in the shallow mantle. Sediments and altered oceanic crust, the top part of the subducted system, are exposed to high temperatures and stresses so they may dehydrate, melt and be scraped off at shallow depths. This component is low in ^8He .

A thin, shallow layer is unlikely to be globally homogenized by plate tectonic processes. The large-scale isotope anomaly in the Southern Hemisphere mantle³⁷ may have been the locus of ancient subduction, just as the circum-Pacific belt is today or the circum-Pangea belt was in the Cenozoic. In this case the Dupal³⁷ anomaly would be due to an ancient subducted component in the shallow mantle (high ^{87}Sr , ^{206}Pb , ^4He). Such an anomaly would be smeared out in time, in the direction of mantle flow, but would diffuse only slowly across the upwellings and downwellings which separate upper mantle convection cells, particularly if it is embedded in buoyant material such as harzburgite.

Note that my model contradicts conventional views of hotspot petrogenesis. The enriched reservoir is shallow rather than deep and it may be relatively infertile, that is, low in CaO and Al_2O_3 . Being shallow it can collect volatiles and LIL by subduction and by trapping of fluids rising from below. Although global it is unlikely to be homogeneous. It protects the deeper, depleted but fertile (low LIL, high CaO and Al_2O_3) reservoir from contamination. It is usually assumed that the source of heat, the source of the basalts and the source of the hotspot geochemical signatures are all the same. The relative stationariness of hot-

spots, the associated swells and high geoid may be due to increased melting in the upper mantle caused by locally high lower mantle temperatures rather than transfer of material from the lower mantle to the surface. It is the thickness of the lithosphere that controls the geochemical signature of intraplate volcanics, not the ease of communication to the lower mantle.

Received 20 May; accepted 28 August 1985.

1. Anderson, Don L. *J. geophys. Res.* **88**, B41-B52 (1983).
2. Menzies, M. & Murthy, V. R. *Nature* **283**, 634 (1980).
3. McCulloch, M. T., Arculus, R., Anappell, B. & Ferguson, J. *Nature* **300**, 166 (1982).
4. Anderson, Don L. *Science* **23**, 89 (1981).
5. Hart, R., Dymond, J. & Hogan, L. *Nature* **278**, 156-159 (1979).
6. Kaneoka, I. *Nature* **302**, 698-700 (1983).
7. Anderson, Don L. *Earth planet. Sci. Lett.* **57**, 1-24 (1982).
8. Chen, C.-Y. & Frey, F. A. *Nature* **302**, 785-789 (1983).
9. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **46**, 323-334 (1980).
10. Dunlop, H. M. & Fitton, J. G. *Contr. Miner. Petrol.* **71**, 125-131 (1979).
11. Barberi, F., Civetta, L. & Varet, J. *Earth planet. Sci. Lett.* **50**, 247-259 (1980).
12. Zhou, X. & Armstrong, R. L. *Earth planet. Sci. Lett.* **58**, 301-329 (1982).
13. Lipman, P. W. & Mehnert, H. H. *Mem. geol. Soc. Am.* **144**, 119-154 (1975).
14. Boettcher, A. L. & O'Neil, J. R. *Am. J. Sci.* **280A**, 594-621 (1980).
15. Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463-513 (1978).
16. Craig, H. & Lupton, J. in *Sea Vol. 1* (ed. Emiliani, C.) 391-428 (Wiley, New York, 1981).
17. Rison, W. & Craig, H. *Earth planet. Sci. Lett.* **66**, 407-426 (1983).
18. Kurz, M., Jenkins, W. J., Hart, S. R. & Clague, D. *Earth planet. Sci. Lett.* **66**, 388-400 (1983).
19. Ozima, M. & Zashu, S. *Science* **219**, 1067-1068 (1983).
20. Hiyagon, H. & Ozima, M. *Earth planet. Sci. Lett.* **58**, 255-264 (1982).
21. Kaneoka, I. *Nature* **302**, 698-700 (1983).
22. Moore, J. G. & Schilling, H. G. *Contr. Miner. Petrol.* **41**, 105 (1973).
23. Stille, P., Unruh, D. M. & Tatsumoto, M. *Nature* **304**, 25-29 (1983).
24. Walck, M. *Geophys. J.R. astr. Soc.* **76**, 697-723 (1984).
25. Nataf, H.-C., Nakanishi, I. & Anderson, Don L. *Geophys. Res. Lett.* **11**, No. 2, 109-112 (1984).
26. Harrison, W. J. *Geochim. cosmochim. Acta* **45**, 1529-1544 (1981).
27. Harrison, W. J. & Wood, B. J. *Contr. Miner. Petrol.* **72**, 145-155 (1980).

The lower mantle, other than being a source of heat, may not be involved at all in petrogenesis.

I thank S. Maaloe, F. Frey and C. Hawkesworth for helpful comments. This work was supported in part by NSF grant EAR811-5236. Contribution 4151, Division of Geological and Planetary Sciences, CalTech, Pasadena.

28. O'Hara, M. *Nature* **243**, 507-508 (1973).
29. Maaloe, S. & Peterson, T. S. *Lithos* **9**, 243-252 (1976).
30. Clague, D. & Bunch, T. E. *J. geophys. Res.* **81**, 4247-4256 (1976).
31. O'Hara, M. J., Saunders, M. J. & Mercy, E. L. P. *Phys. Chem. Earth* **2**, 571-604 (1975).
32. Bender, J. F., Hodges, F. N. & Bence, A. E. *Earth planet. Sci. Lett.* **41**, 277-302 (1978).
33. Frey, F. A., Bryan, W. B. & Thompson, G. J. *geophys. Res.* **79**, 5507-5527 (1974).
34. O'Nions, R. K., Pankhurst, R. J. & Gronvold, K. J. *Petrol.* **17**, 315-338 (1976).
35. Leeman, W. P., Murali, A. V., Ma, M.-S. & Schmitt, R. A. *Oregon Dept Geol. Mineral Indust. Bull.* **96**, 169-183 (1977).
36. de Argollo, R. M. & Schilling, J. G. *Nature* **276**, 24-28 (1978).
37. Hart, S. R. *Nature* **309**, 753-757 (1984).
38. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *Earth planet. Sci. Lett.* **34**, 13-22 (1977).
39. Carlson, R. W., Lugmair, G. W. & Macdougall, J. D. *Geochim. cosmochim. Acta* **45**, 2483-2499 (1981).
40. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 5, 249-252 (1976).
41. Zindler, A., Jagoutz, E. & Goldstein, S. *Nature* **298**, 519-523 (1982).
42. Hawkesworth, C. J., Rogers, N. W., van Calsteren, P. W. & Menzies, M. A. *Nature* **311**, 331-335 (1984).
43. McCulloch, M. T., Jaques, A. L., Nelson, D. R. & Lewis, J. D. *Nature* **302**, 400-403 (1983).
44. Vollmer, R. & Norry, M. J. *Nature* **301**, 141-143 (1983).
45. Schilling, J. G. *Nature* **242**, 565-571 (1973).
46. Sun, S.-S. *Phil. Trans. R. Soc. A297*, 409-445 (1980).
47. Oversby, V. M. & Gast, P. W. *Earth planet. Sci. Lett.* **5**, 199-206 (1968).
48. Oversby, V. M. & Gast, P. W. *J. geophys. Res.* **75**, 2087 (1970).
49. Doe, B. R., Leeman, W. P., Christiansen, R. L. & Hedge, C. E. *J. geophys. Res.* **87**, 4785-4806 (1982).
50. Kaneoka, I., Takaoka, N. & Clague, D. A. *Earth planet. Sci. Lett.* **66**, 427-437 (1983).
51. Kurz, M. D., Jenkins, W. J. & Hart, S. R. *Nature* **297**, 43-47 (1982).
52. Hart, R., Dymond, J., Hogen, L. & Schilling, J. G. *Nature* **305**, 403-407 (1983).
53. Kurz, M. D. & Jenkins, W. J. *Earth planet. Sci. Lett.* **53**, 41-54 (1981).

Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors

Gordon Keller*, Christopher Paige*, Eli Gilboa† & Erwin F. Wagner‡

* Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland

† Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA

‡ European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, FRG

Bone marrow cells infected with retroviral vectors carrying the bacterial neomycin resistance (neo) gene as a marker were used for long-term reconstitution of the haematopoietic system of irradiated mice. The neo gene is expressed in the myeloid and lymphoid lineages of these animals and an analysis of the sites of viral integration indicates that these lineages are derived from the same primitive multipotent cells.

THE transfer of foreign genes with retroviral vectors to precursors of the haematopoietic system provides a unique opportunity to analyse the expression and function of a variety of different genes in a well-characterized developmental system that can be manipulated both *in vitro* and *in vivo*. In addition, the proviral integration site can be used as a clonal marker for identifying haematopoietic precursors and for following their progeny through differentiation. This is of great interest as our knowledge of the structure of the haematopoietic system, particularly the relationship of the lymphoid lineages to one another or to the other (myeloid) blood cell lineages, is still incomplete. Colony assays have defined a variety of different precursors, including those which are pluripotent, oligopotent and unipotent for the myeloid lineages¹⁻⁵. However, demonstration of the more primitive precursors capable of generating both lymphoid and myeloid progeny, as well as of precursors restricted to the lymphoid lineages, has been more difficult.

Unique chromosome marker experiments⁶⁻⁸ and embryo reconstitution studies⁹ in the mouse as well as isoenzyme analysis in patients suffering from chronic myelogenous leukaemia

(CML)¹⁰ all support the concept of a common precursor for the lymphoid and myeloid lineages. At present, only one report has provided evidence for the existence of long-lived precursors restricted to the T-cell lineage⁶. Precursors committed to the B-cell lineage or lymphoid precursors common to both the B- and T-cell lineages have not yet been described.

Retroviral vectors will be useful for both gene expression and lineage analysis studies, provided that they can infect all classes of precursors and that the transferred genes can be expressed in the progeny of the infected cells. In the present report, retroviral vectors were used to transfer the neomycin resistance (*neo*) gene to a variety of haematopoietic precursors, including those capable of participating in long-term reconstitution of irradiated mice. Expression of the introduced *neo* gene was demonstrated in the different blood cell lineages of all recipients analysed. Unique viral integration sites in the DNA of different tissues, purified cell populations and cloned T- and B-cell hybridomas showed that multipotent precursors capable of generating both lymphoid and myeloid progeny were infected with the recombinant vectors.

Infection of bone marrow cells

In order to achieve a high efficiency of retrovirus infection, the protocol shown in Fig. 1 was developed. The structure of the retroviral vectors (N2 and N3) used for infection is shown in Fig. 2a. N2 was generated helper-free, whereas N3 contains helper virus (Moloney murine leukaemia virus, M-MuLV). The essential features of the protocol include a 24-h co-culture of the bone marrow cells with the virus-producing cells, followed by a 48-h preselection step that enriches for precursors expressing the *neo* gene. In three separate experiments 50–95% of all colony-forming cells (CFC) survived the 24-h co-culture with the virus-producing cells and of these, 10–20% were able to grow and form colonies in the presence of G418. The frequency of G418-resistant (G418^r) CFC was similar following co-culture with cells producing either N2 or N3 virus. Following the preselection step in liquid culture, 15–50% of the input number

of CFC were recovered. However, a high percentage (60–95%) of these were G418^r.

Reconstitution and analysis

Virus-infected cells from unselected and preselected populations were used to reconstitute irradiated recipients. Myeloid and lymphoid tissues from these mice were analysed at different

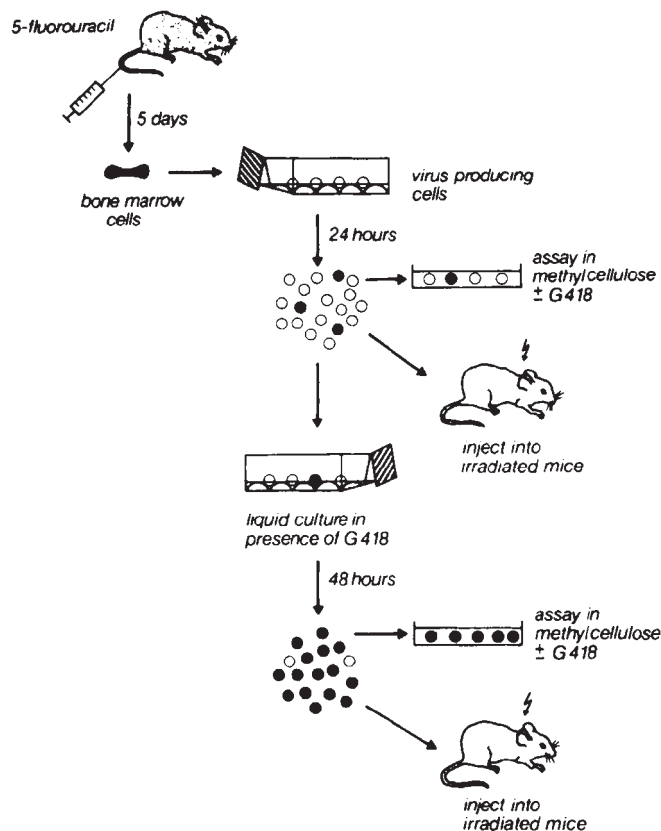


Fig. 1 Protocol for viral infection of bone marrow cells. Bone marrow cells from mice treated 5 days earlier with 5-fluorouracil (150 mg per kg body weight) were seeded directly onto a monolayer of irradiated (1,000 R X-ray) virus-producing cells. CBA-HT6T6Bii mice (>10 weeks old), obtained from the Basel Institute for Immunology breeding colony, were used as a source of donor bone marrow in all experiments. The virus-producing cells were maintained in IMDM with 5% FCS and subcultured the day before infection, to achieve 50–80% confluency on the day of the experiment. For co-culture, half of the supernatant was removed and replaced with fresh medium containing the bone marrow cells, polybrene ($10 \mu\text{g ml}^{-1}$), conditioned medium from WEHI-3B (D⁺) cells (20%), transferrin ($30 \mu\text{g ml}^{-1}$) and BSA (1 mg ml^{-1}). For reconstituted mice 1, 2, 3 and 4, horse serum (15%) and hydrocortisone (10^{-7} M) were also included, whereas for mouse 5 these two reagents were replaced by medium conditioned by adherent bone marrow cells. Either 25-cm² flasks containing $2-3 \times 10^6$ bone marrow cells in a final volume of 4 ml of medium or 75-cm² flasks containing $5-10 \times 10^6$ cells in 10 ml of medium, were used. After 24 h of co-culture, the bone marrow cells were collected. A sample from this infected population was assayed in methylcellulose cultures to determine the frequency of G418^r CFC (see Table 1). A second sample was injected into irradiated (900R) recipients (CBA/N; >12 weeks old). The remaining cells were selected in G418 (2 mg ml^{-1}) for 48 h before assaying for (1) G418^r CFC and (2) the ability to reconstitute irradiated recipients. This preselection step was carried out on the virus-producing cells for mice 1, 2, 3 and 4, and in the absence of any adherent layer for mouse 5. WEHI-3B (D⁺) conditioned medium was included in the preselection cultures.

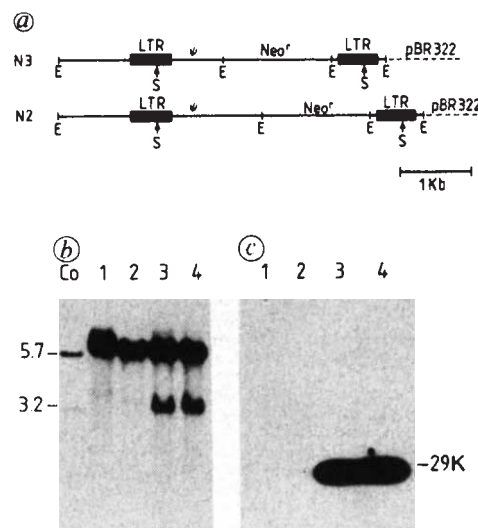


Fig. 2 a, Structure of the *neo*-containing retroviruses. The two constructs (N3 and N2) are composed of three parts: (i) the 5'-long terminal repeat (LTR) fragment containing the viral packaging signal (ψ) was derived from M-MuLV small circular DNA by excision with *Bgl*I and *Xho*I, limited exonuclease digestion with *Bal*31 and addition of *Eco*RI linkers (more viral sequences are present downstream from the 5'-LTR in N2 than in N3); (ii) the bacterial *neo* gene (*Neo*^r) derived from Tn5 by excision with *Bgl*II and *Bam*HI and addition of *Eco*RI linkers; and (iii) a 3'-LTR fragment which was also derived from M-MuLV small circular DNA by excision with *Bam*HI and *Pst*I, followed by *Bal*31 digestion, addition of *Eco*RI linkers and digestion with *Cl*aI. The DNA constructs were generated by standard recombinant DNA techniques, and introduced into a thymidine kinase-negative (TK⁻) derivative of NIH 3T3 cells (N3 virus) or into ψ 2 cells²⁴ (N2 virus) by the calcium phosphate precipitation method. NIH 3T3 or ψ 2 cells were selected with G418 (1 mg ml^{-1}) for 8–10 days. NIH 3T3 cells carrying the N3 provirus were subsequently superinfected with M-MuLV helper virus and supernatants from these cells were harvested and used to infect fresh NIH 3T3 cells. Following G418 selection, single clones were isolated, expanded, and used as a source of recombinant virus. N2-transfected ψ 2 cells were also selected in G418 (1 mg ml^{-1}) and single resistant clones were expanded to produce helper-free virus. N3 was produced at a titre of 10^6 G418^r colony forming units CFU per ml together with $1-2 \times 10^6$ XC units of helper virus. N2 was produced helper-free at 5×10^5 G418^r CFU ml⁻¹. Supernatants from NIH 3T3 cells infected with N2 virus did not contain G418^r CFU (determined by infection of fresh NIH 3T3 cells). When cells were infected with helper M-MuLV, the supernatant contained 10^5 G418^r CFU ml⁻¹. This assay can detect the presence of one infectious M-MuLV particle in the N2 virus preparation. b, Analysis of intact proviral DNA sequences in spleen cells from mice that had been reconstituted for 12 days with unselected (equivalent of 850 mix-CFC, 15% G418^r) or preselected (equivalent to 170 mix-CFC, >80% G418^r) cells. DNA of high *M*_r was isolated from the spleens of four reconstituted animals and digested with *Sac*I that cuts once in each of the LTRs of the *neo*-containing and helper proviruses and once within the helper virus genome. Equal amounts ($10 \mu\text{g}$) of digested DNA were fractionated on agarose gels, transferred to nitrocellulose by the method of Southern¹², and hybridized with a ³²P-labelled U3-specific probe²⁵. This probe detects the diagnostic 3.2-kb fragment of the intact *neo* provirus and a characteristic 5.7-kb fragment from the helper virus genome. The efficiency of bone marrow infection can be estimated from the band intensities by comparison with the cell line Co, which carries only a single copy of the *neo* gene ($2 \mu\text{g}$ of DNA was loaded in this lane). Spleen DNAs of two mice, reconstituted with unselected cells, are shown in lanes 1 and 2. Lanes 3, 4 are DNAs from those mice, reconstituted with preselected bone marrow cells. c, NPT-II activity from spleen extracts of the four mice described in b. Prior to DNA extraction, the frozen cell pellets ($\sim 10^7$ cells) were thawed and lysed in $200 \mu\text{l}$ of H₂O. The suspension was centrifuged for 5 min at 10,000 r.p.m. and the aqueous supernatant was used for the NPT-II assay¹¹. Equal amounts of soluble protein were separated in a non-denaturing polyacrylamide gel, which was then overlaid with a 1% agarose gel containing kanamycin sulphate ($100 \mu\text{g ml}^{-1}$) and [γ -³²P]ATP ($1.5 \mu\text{Ci ml}^{-1}$). The labelled kanamycin sulphate was transferred onto P81 paper which was washed three times (5 min) with hot water (80°C), dried and exposed on X-ray film. The labelled 29,000-*M*_r band is diagnostic for the NPT-II activity.

times after reconstitution for expression of the *neo* gene. Expression was demonstrated by the presence of G418^r CFC and by the presence of the *neo* gene product, as determined by the neomycin phosphotransferase (NPT-II) assay¹¹. The presence of the intact proviral structure as well as the existence of unique viral integration sites in these mice was analysed in DNA from the different organs (bone marrow, thymus, lymph node, spleen), from purified cell populations and from hybridomas, by Southern blotting¹².

Table 1 Bone marrow cells used for reconstitution of irradiated mice

Mouse	Pre-selected with G418	Total cell no. injected	No. of mix-CFC injected	% G418 ^r mix-CFC
1	+	2.1 × 10 ⁶	1,350	86
1*	—	5.0 × 10 ⁶	750	26
2	+	2.2 × 10 ⁶	8,100	88
2*	—	5.0 × 10 ⁶	2,200	36
3	—	1.0 × 10 ⁶	2,950	25
4	—	1.0 × 10 ⁶	2,950	25
5	+	4.2 × 10 ⁶	23,100	65
5*	—	5.0 × 10 ⁶	1,150	6
5**	—	5.0 × 10 ⁶	1,170	1

Mice 1–4 were reconstituted with bone marrow cells infected with N3 virus whereas mouse 5 received N2(helper-free)-infected bone marrow cells. No viraemia was detected in the sera of mice 1–4 by a sensitive reverse transcriptase assay despite the fact that competent helper virus was present. Mixed colonies were grown from bone marrow cells in 1-ml cultures (35-mm Petri dishes) containing 0.9% methylcellulose in Iscove's modified Dulbecco's medium (IMDM), 5% fetal calf serum (FCS), 10 mg de-lipidated and deionized bovine serum albumin (BSA), 300 µg iron-saturated transferrin, 22.4 µg dipalmitoyl phosphatidylcholine, 23.5 µg cholesterol and 17.3 µg oleic acid. All reagents were prepared as described by Iscove²⁰. DEAE-Sepharose-purified WEHI-3B(D⁺) conditioned medium as a source of multilineage haematopoietic growth factor (MultiHGF/BPA/IL-3)¹⁴ and serum from anaemic mice as a source of erythropoietin²¹ were added in sufficient amounts to give optimal growth of mixed colonies. Conditioned medium from plastic-adherent bone marrow cells was added (10%) as a source of additional growth factors²². G418 was included at 2 mg ml⁻¹ as indicated. Colonies were counted after 9–12 days at 37 °C in a humidified environment of 5% CO₂ in air. Mixed colony-forming cells (mix-CFC) are those precursors that generate colonies containing cells from the erythroid plus at least two other lineages. *Secondary recipient of bone marrow from indicated mice. **Tertiary recipient. This animal was not healthy at the time of killing. The thymus and lymph nodes were very small and the marrow was somewhat hypocellular. This is not a surprising finding, as it is well known that the reconstituting capacity of bone marrow decreases dramatically with each passage through an irradiated recipient²³.

Short-term reconstituted mice

The efficiency of infection and preselection was analysed in mice reconstituted for 12 days with N3-infected bone marrow cells. Spleens from both groups contained large numbers of macroscopic colonies. Methylcellulose cultures showed a higher frequency of G418^r CFC in the spleens of mice reconstituted with preselected cells (70%) compared with those that received the unselected population (5%).

To analyse the efficiency of gene transfer and the structure of the provirus, DNAs were prepared from individual spleens of both groups. Southern analysis showed that spleens of mice reconstituted with preselected cells contained more intact *neo*-specific sequences (3.2 kilobases, kb) than DNA from spleens of mice reconstituted with unselected cells (Fig. 2b). Similar amounts of helper virus DNA were found in both groups, although no replicating helper virus was detectable by a reverse transcriptase assay. It is possible that CBA/N mice inhibit viral replication, as others have observed a similar phenomenon (W. Ostertag, personal communication). When cell lysates were tested for an active *neo* gene product, much higher levels of NPT-II activity were found in spleen extracts of mice reconstituted with preselected cells compared with those reconstituted with the unselected population (Fig. 2c).

These findings establish that precursors capable of short-term reconstitution can be infected with the recombinant *neo*-containing virus and can survive the preselection step. In order to establish that primitive precursors capable of long-term reconstitution had been infected with the *neo*-containing virus, we next analysed mice reconstituted for 6–17 weeks. The self-renewal capacity of these precursors was assessed by transferring bone marrow from the primary reconstituted animals (mice 1, 2 and 5) to secondary recipients (1*, 2* and 5*) and in one case to a tertiary recipient (5**). Table 1 lists the different bone marrow populations used for reconstitution.

Long-term reconstituted mice

Expression of the *neo* gene was analysed first in the myeloid lineages by the ability of bone marrow cells from the reconstituted animals to form colonies in methylcellulose cultures containing G418. Independent of the time after reconstitution, all of the animals assayed contained G418^r CFC (Table 2). Those reconstituted with preselected cells had higher frequencies of G418^r CFC (13–37%) than those which received unselected cells (6% and 1% for mice 3 and 4). These findings indicate that the *neo* gene is being expressed in precursors of the myeloid lineages. It is noteworthy that the frequency of G418^r CFC decreased with time in the primary recipient and after transfer of bone marrow to a secondary host (Table 2).

To determine whether or not the *neo* gene was present and expressed in cells of the B-lymphoid lineage, clonable B cells

Table 2 G418-resistant colony-forming cells in reconstituted mice

Mouse	Time of reconstitution (weeks)	Colonies per 10 ⁶ cells						% Total G418 ^r CFC
		-G418			+G418			
		E	Mix	Total	E	Mix	Total	
1	7	220	150	2,470	60	40	780	32
1*	3	—	—	880	—	—	180	21
2	6	490	420	4,150	110	150	1,530	37
2*	4	53	40	963	—	—	100	10
3	9	450	490	3,570	30	20	200	6
4	17	240	470	2,470	2	4	20	1
5	11	310	360	2,370	34	21	305	13
5*	5	125	230	1,425	2	3	19	1
5**	7	130	80	1,900	0	0	3	0.2

Conditions for growth of colonies as described for Table 1. Colonies were classified as follows: E, colonies in which all or nearly all of the cells are erythroid; Mix, mixed colonies containing cells from the erythroid plus at least two other lineages; Total, all colonies in the culture, including E, mix, plus those comprised of mast cells, neutrophils, macrophages and mixtures of these. There were no detectable differences in the appearance of colonies or in the state of maturation of the cells within the colonies in the presence or absence of G418 (2 mg ml⁻¹). The mean numbers of colonies counted in 2–6 replicate cultures are shown. Most long-term reconstituted mice had normal numbers of bone marrow cells (16–36 × 10⁶ cells per two femurs and two tibias). The only exceptions were mice 1* and 5**, which contained 3.0 × 10⁶ and 12 × 10⁶ bone marrow cells respectively.

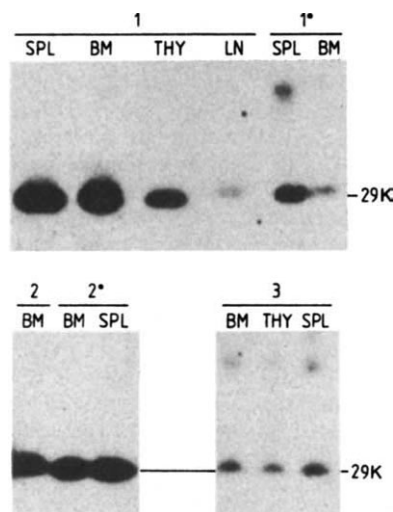


Fig. 3 Neomycin-phosphotransferase (NPT-II) activity in whole tissue extracts from long-term reconstituted mice. Preparation of tissue extracts and the *in situ* NPT-II assay are described in Fig. 2c. SPL, spleen; BM, bone marrow; THY, thymus; LN, lymph node; approximately five times less protein was loaded onto the LN lane. Animals 1 and 2 were reconstituted with preselected bone marrow cells, whereas mouse 3 received unselected cells. Mice 1* and 2* are recipients of bone marrow from mice 1 and 2 respectively.

(CFU-B), a subpopulation of mature B cells that are able to form colonies *in vitro*¹³, were grown in the presence of G418. Between 1 and 3% of the CFU-B found in the spleens of mice 2, 3, 4 and 5 were G418^r. Only 0.3% of those from mice 1* and 5* were able to grow in the presence of G418. Antibody secretion from cells in these G418^r colonies was demonstrated by their ability to form plaques as monitored by protein A plaque assay¹³.

In addition to growth of cells in G418, *neo* gene expression was monitored by assaying different tissue extracts from the reconstituted mice for NPT-II activity. Both myeloid and lymphoid organs of all long-term reconstituted mice tested expressed NPT-II activity (Fig. 3). As expected, organs from mice reconstituted with preselected cells (mice 1 and 2) contained higher enzyme levels than those from mice reconstituted with unselected cells (mouse 3). The presence of the *neo*-encoded protein in the different organs, together with the ability of the various precursors to grow in cultures containing G418, demonstrates unequivocally that the introduced *neo* gene is being expressed in all haematopoietic lineages.

Cell lineage analysis

The next series of experiments was designed to determine the number and the pattern of viral integration sites in the organs and in individual cell lineages of the reconstituted animals. Assuming that the site of integration of the provirus is random, information obtained from this analysis allows (1) estimation of the number of infected precursors that have participated in reconstitution of the different organs and (2) determination of the differentiation potential and longevity of these precursors.

Mice 1, 1*, 2, 2* and 3. Southern analysis of *Bam*HI-digested DNA from thymus, lymph node, spleen and bone marrow of mouse 1 showed three discrete fragments of 6.8, 9.4 and 16 kb (Fig. 4). These fragments were also present in bone marrow cells from the secondary recipient 1*. This finding suggests that a maximum of three precursors gave rise to the majority of *neo*-containing cells in the different organs of these mice. The intensity of the three bands compared with the intensity of a gene present as a single copy in all cells (hybridoma lanes in Fig. 4) indicates that a high proportion of the cells contain the *neo* provirus. From the presence of the same three bands in the different tissues, we conclude that the infected precursors were multipotent with the capacity to generate cells of the myeloid and lymphoid lineages. Finally, the presence of the same three integration sites in bone marrow of mouse 1*, indicates that this precursor (or precursors) could generate new cells capable of

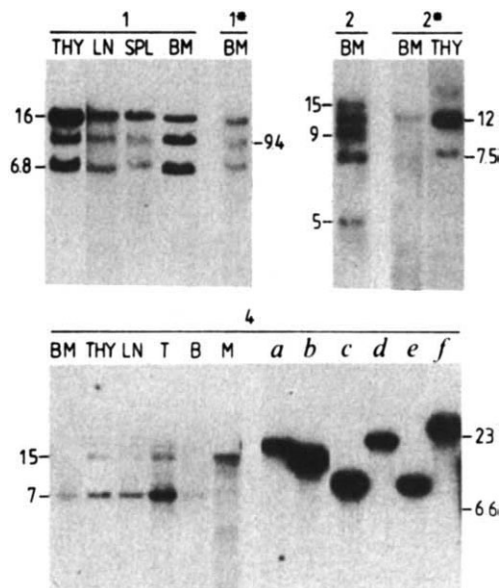


Fig. 4 Lineage analysis of three long-term reconstituted mice (1, 2 and 4). High-*M_r* DNA was isolated from various tissues, purified cell populations and from B-cell hybridomas and digested with *Bam*HI, which does not cut the proviral genome. Approximately 10 µg of digested DNA was fractionated on agarose gels and analysed by Southern blotting¹² using a *neo*-specific probe. 1* and 2* are recipients of bone marrow from mice 1 and 2. BM, bone marrow; THY, thymus; LN, lymph node; SPL, spleen; T, IL-2-dependent spleen cells; B, surface immunoglobulin-positive (sIg⁺) spleen cells; M, mast cells; a-f, B-cell hybridomas.

Methods. To derive IL-2-dependent cells, spleen cells were first stimulated with concanavalin A (Con A, 48 h) then cultured in medium containing either recombinant IL-2 (Roche) or supernatant from EL-4 cells²⁶ for 1-2 weeks. Over 98% of the cells used for DNA analysis were Thy-1.2-positive as determined by direct immunofluorescence using a labelled monoclonal antibody anti-Thy-1.2 and a fluorescence activated cell sorter (FACS II) for analysis. sIg⁺ cells were obtained following labelling of spleen cells with a fluorescein isothiocyanate-conjugated sheep anti-mouse antibody and sorting on a FACS II. Over 95% of the cells in the sorted population were sIg⁺. Mast cells^{27,28} were derived from bone marrow cells that were grown for more than 3 weeks in the presence of WEHI-3B(D⁻) conditioned medium and G418 (2 mg ml⁻¹). Mast cells showed the characteristic granules for mast cells determined by astra blue²⁹ and toluidine blue staining³⁰. B-cell hybridomas were generated by fusion between the nonproducer F0 line and lipopolysaccharide (LPS)-stimulated B-cell blasts. Fusions were carried out 4 days after initiation of *in vitro* cultures with 50 µg ml⁻¹ lipopolysaccharide (LPS). G418 (1-2 mg ml⁻¹) was included on day 2 of these cultures. The LPS blasts were washed and fused according to methods described elsewhere³¹. Established hybridomas were selected for resistance to G418 before DNA analysis.

participating in the reconstitution of a second irradiated recipient. When DNA from the spleen of mouse 1 was digested with a second enzyme (*Hind*III), only three discrete bands were again detected, strongly suggesting that no more than three integration sites were present (data not shown).

DNA from bone marrow of mouse 2 contained at least five distinct integration sites (Fig. 4). On transfer of the bone marrow to mouse 2*, three of the hybridizing bands (5, 9 and 15 kb) were undetectable. Two of the bands found in the DNA of the primary bone marrow could also be found in the DNA from the thymus of mouse 2*. In addition, a fragment of higher relative molecular mass (*M_r*), which was not found in the primary mouse, was also present in the thymus DNA. Only the 12-kb band found in the bone marrow of mouse 2 was present in the DNA of bone marrow cells from mouse 2*. As this 12-kb fragment is also found in the thymus DNA of mouse 2*, it could represent a clone of cells derived from a primitive multipotent precursor. Again, the intensity of the bands in the bone marrow of mouse 2 and the thymus of mouse 2* strongly suggests that the majority of cells in these tissues contain the *neo* gene.

Mouse 3, reconstituted with unselected cells, contained a single integration site in the DNA from bone marrow, thymus, spleen and lymph node (data not shown). Only a small proportion of the cells in this animal had the *neo* gene, as determined by band intensity and colony growth (see Table 2). These

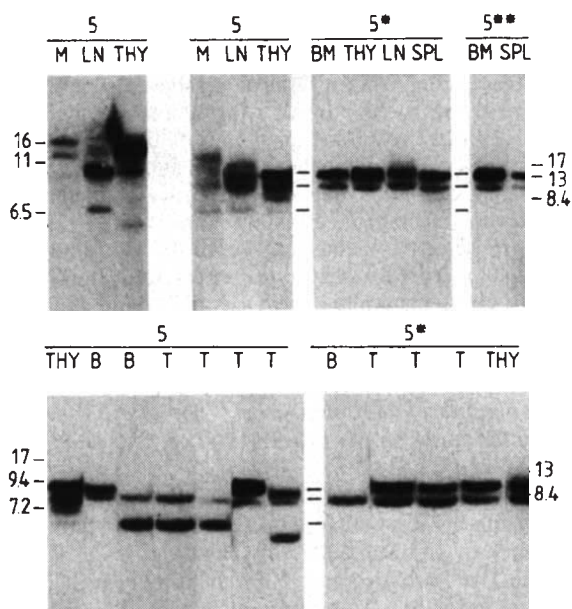


Fig. 5 Lineage analysis of mice 5, 5* and 5**. Southern analysis was carried out as described in Fig. 4 legend. DNA from all tissues, cell populations and hybridomas was digested with *Hind*III, an enzyme that does not cut the provirus genome. In addition, DNA from tissues of mouse 5 was also analysed following *Bam*HI digestion (see upper left-hand corner). Tissues and cell populations are designated as in Fig. 4 except for T, which refers to T-cell hybridomas.

Methods. T-cell hybridomas were generated by fusion between the thymoma BW5147 and spleen cells that had been stimulated with Con A in either the presence or absence of G418. Fusion was accomplished by standard techniques using polyethylene glycol 4000 (Merck) as a fusing agent. The hybridomas were first selected in hypoxanthine/aminopterin/thymidine medium and then in G418 (at least 1 mg ml^{-1} G418). From mouse 5, 8×10^6 cells stimulated without G418 and 3.5×10^6 cells stimulated in the presence of G418 were separately fused with a fivefold excess of BW5147 cells, then distributed into 24-well plates (3×10^5 cells per well). After several weeks all the wells from all the plates contained hybridomas; over 80% of the wells had cells that grew in G418. These G418^r cells were expanded and cloned by limiting dilution. From mouse 5*, 8×10^5 Con A blasts that had been stimulated in the presence of G418 were fused to 4×10^6 BW5147 cells and then seeded into two 24-well plates. This fusion generated six hybridomas, all of which were G418^r.

findings raised the question of whether the gene is present in cells of a number of different lineages or in only a single lineage that is found in all the organs examined. This problem was addressed in the analysis of mouse 4 by expanding cells of specific lineages *in vitro* in the presence or absence of G418 and including these in the DNA analysis.

Mouse 4. Interleukin-2 (IL-2)-dependent cells (T cells), and surface immunoglobulin-positive cells derived from the spleen and B-cell hybridomas were used to analyse the integration sites in the lymphoid lineages. Multilineage haematopoietic growth factor (MultiHGF/BPA/IL-3)¹⁴-dependent mast cells represented the myeloid lineages. A common band of 15 kb was found in the DNA from the thymus, lymph node, IL-2-dependent cells, mast cells and one of the six B-cell hybridomas (Fig. 4); this shows that cells from distinct lineages carry the same integration site and indicates that cells present in each of these populations arose from a common multipotent precursor. A similar interpretation is valid for the 7-kb band which is also present in the myeloid and lymphoid lineages.

Despite the fact that the organ DNAs showed two major integration sites, the cloned B-cell hybridomas exhibited five independent sites (Fig. 5). Only one of the fragments found in the B-cell hybridomas corresponded to the predominant 15-kb fragment found in the DNA from the tissues. Several bands found in other hybridomas corresponded to fainter bands that are visible in the DNA of the tissues, the IL-2-dependent cells and the mast cells. However, at least two could not be detected in the DNA of any cell lines or organs (Fig. 4d,f). With this analysis, it is not clear whether the unique fragments in the B-cell hybridomas are also present at low levels in the DNA of

the other lineages or whether they are restricted to the B-cell lineage.

To confirm that cells immortalized as hybridomas were of the B-cell lineage, DNA was next analysed for immunoglobulin heavy-chain gene rearrangements. Each hybridoma contained rearranged alleles and was found to be clonal in this analysis (data not shown).

Mice 5, 5* and 5.** The most detailed lineage analysis was carried out on mice reconstituted with N2-infected bone marrow cells. DNA from the organs, from mast cells and from B- and T-cell hybridomas of mouse 5 was analysed for the presence of unique viral integration sites. A similar analysis was performed with DNA from the organs, from T- and B-cell hybridomas of mouse 5*, and with DNA from the bone marrow and spleen of mouse 5**.

Two common integration sites of 17 and 13 kb were found in *Hind*III-digested DNA from all organs of the three mice (Fig. 5). The intensity of the hybridizing bands found in each of the tissues compared with the intensity of the bands found in the hybridomas suggests that a high proportion of cells are carrying the *neo* gene and are therefore of donor origin. This finding was confirmed by the presence of the T6 chromosome in more than 90% of metaphases from bone marrow cells of mice 5* and 5**. In addition to the two common bands, a fragment of lower M_r (8.4 kb) was found in mast-cell, lymph-node and thymus DNA of mouse 5. This fragment was present at low amounts in DNA from the thymus and spleen of mouse 5*; however, it was not detected in DNA from the spleen and bone marrow of the tertiary recipient (5**). Another integration site of 9.4 kb was found in the thymus of mouse 5, and in the spleen of mouse 5*. Finally, DNA from the mast cells and lymph node of mouse 5 contained fragments of higher M_r that were not detected in any other tissue. When DNA from mast cells, lymph node and thymus of mouse 5 was digested with *Bam*HI, at least four integration sites were found in each. Two of these sites (16 and 11 kb) appear to be common to all tissues and may correspond to the common sites present in the *Hind*III-digested DNA. The additional sites may be restricted to individual lineages.

The initial analysis of a limited number of B- and T-cell hybridomas from mice 5 and 5* is shown in the lower panel of Fig. 5. The two common bands present in the tissues are also found in one B-cell and two T-cell hybridomas from mouse 5 as well as in three T-cell hybridomas from mouse 5*. One of these T-cell hybridomas of mouse 5 contained, in addition, a third integration site of 7.2 kb, which was not detected in any of the tissues. Three other hybridomas of mouse 5 (one B-cell and two T-cell) also contained two sites that seem to be present in the tissues from this mouse. One B-cell hybridoma from mouse 5* had only one of the two common sites. This finding is difficult to interpret without further analysis. A more detailed study of the nature of the integration sites in the B- and T-cell hybridomas of these mice will be published elsewhere.

The presence of the same integration sites in mast cells, in B- and T-cell hybridomas, as well as in different tissues, is the most convincing demonstration to date of the existence of multipotent precursors able to generate myeloid and lymphoid cells. The persistence of integration sites in transfer from the primary to the secondary and tertiary recipients, suggests that the multipotent cell infected originally *in vitro*, exhibited substantial self-renewal capacity.

Discussion

The present report describes an efficient protocol for introducing retroviral vectors carrying a foreign gene into haematopoietic precursors of the mouse as well as the application of this method to cell lineage analysis. The protocol differs from methods used by others^{15,16} in several ways. First, bone marrow from 5-fluorouracil-treated rather than from normal mice was used for infection as it is enriched for primitive precursors¹⁷. Second, a preselection step was included in the protocol to obtain a cell population with a high frequency of infected precursors that were expressing the *neo* gene.

Using this protocol we have demonstrated that multipotent precursors common to the myeloid and lymphoid lineages can be infected with recombinant retroviruses. Furthermore, we have shown by the NPT-II assay and by growth of precursors in G418 that the introduced gene is expressed in all lineages of all reconstituted animals tested. Several other groups have recently described the transfer of genes to bone marrow cells using retroviral vectors^{15,16,18}. However, no-one has shown the infection of primitive multipotent precursors or expression of an introduced gene in distinct myeloid and lymphoid lineages.

From analysis of proviral integration sites we were able to enumerate the numbers of infected precursors which reconstituted each recipient. Through transfer of bone marrow from one recipient to the next, it was possible to estimate the longevity (self-renewal capacity) of such cells. Results from mice 1*, 2*, 5* and 5** indicated that some clones were maintained through transplantation, while others were lost (Figs 4, 5). Those clones that were lost might represent precursors at an intermediate stage of development while those that persisted could have been derived from the most primitive ones exhibiting self-renewal capacity. At least six multipotent precursors were identified in five different mice and four of these demonstrated self-renewal ability (Figs 4, 5). Of equal interest would be the identification of restricted precursors and in particular those restricted to the lymphoid lineages. Clones representing cells that might be restricted to the T-cell lineage are those marked by the 6.5-kb

fragment in the thymus and the 7.2-kb band in one T-cell hybridoma of mouse 5 (Fig. 5). Further analysis of viral integration sites from more reconstituted mice should identify such restricted precursors, thereby allowing a novel approach to studying cell lineage relationships.

In conclusion, we have shown that a foreign gene carried by a recombinant retrovirus can be efficiently introduced into haematopoietic precursors and that this gene is expressed in the progeny of these cells. We have also used the viral integration site as a clonal marker for identifying, enumerating and following the fate of haematopoietic precursors. Through analysis of unique viral integration sites, we have demonstrated unequivocally that cells with distinct lymphoid and myeloid phenotypes arose from a common precursor.

Since submission of this manuscript, Dick *et al.*¹⁹ have published a similar study on the introduction of a selectable gene into haematopoietic cells.

We thank Karin Damlin and Helena Skarvall for technical assistance, William Leiserson for FACS analysis and Dr Michael Julius for helpful advice and for providing IL-2. We also acknowledge Drs Barbara Fagg, Norman Iscove, Rolf Müller, Colin Stewart and Ulrich Rüther for critically reading the manuscript, Marion Kennedy for help in its preparation, and Lynn Gisler and Ines Benner for typing it. E.G. is a recipient of a grant from ACS no. MV116B. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche Co., Ltd., CH-4005 Basel, Switzerland.

Received 25 July; accepted 12 September 1985

1. Till, J. E. & McCulloch, E. A. *Radiat. Res.* **14**, 213-222 (1961).
2. Metcalf, D. *Recent Results in Cancer Research* (Springer, Berlin, 1977).
3. Johnson, G. R. & Metcalf, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3879-3882 (1977).
4. Iscove, N. N. in *Haematopoietic Cell Differentiation* (eds Golde, D. W., Cline, M. J., Metcalf, D. & Fox, C. F.) 37-52 (Academic, New York, 1978).
5. Suda, T., Suda, J. & Ogawa, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6689-6693 (1983).
6. Wu, A. M., Till, J. E., Siminovich, L. & McCulloch, E. A. *J. exp. Med.* **127**, 455-463 (1968).
7. Edwards, G. E., Miller, R. G. & Phillips, R. A. *J. Immun.* **105**, 719-729 (1970).
8. Abramson, S., Miller, R. G. & Phillips, R. A. *J. exp. Med.* **145**, 1567-1579 (1977).
9. Mintz, B., Anthony, K. & Litwin, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7835-7839 (1984).
10. Fialkow, P. J. *J. cell. Physiol. Suppl.* **1**, 37-43 (1982).
11. Reiss, B., Sprengel, R., Will, H. & Schaller, H. *Gene* **30**, 211-218 (1984).
12. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
13. Paige, C. J. & Skarvall, H. *J. immun. Meth.* **52**, 51-61 (1982).
14. Iscove, N. N. in *Leukemia: Dahlem Konferenzen* (ed. Weissman, I. L.) (Springer, Berlin, in the press).
15. Williams, D. A., Lemischka, I. R., Nathan, D. G. & Mulligan, R. C. *Nature* **310**, 476-480 (1984).

16. Miller, A. D., Eckner, R. J., Jolly, D. J., Friedmann, T. & Verma, I. M. *Science* **225**, 630-632 (1984).
17. Hodgson, G. S. & Bradley, T. R. *Nature* **281**, 381-382 (1979).
18. Joyner, A., Keller, G., Phillips, R. A. & Bernstein, A. *Nature* **305**, 556-558 (1983).
19. Dick, J. E., Magli, M. C., Huszar, D., Phillips, R. A. & Bernstein, A. *Cell* **42**, 71-79 (1985).
20. Iscove, N. N. in *Methods for Serum-Free Cultures of Neuronal and Lymphoid Cells* (Liss, New York, 1984).
21. Tambourin, P. E., Wendling, F., Gallien-Lartigue, O. & Huauclme, D. *Biomedicine* **19**, 112-116 (1973).
22. Iscove, N. N., Keller, G. & Roitech, C. in *Stem Cell Physiology* (ed. Palek, J.) (Liss, New York, 1985).
23. Harrison, D. E. & Astle, C. M. *J. exp. Med.* **156**, 1767-1779 (1982).
24. Mann, R., Mulligan, R. C. & Baltimore, D. *Cell* **33**, 153-159 (1983).
25. Müller, R. & Müller, D. *EMBO J.* **3**, 1121-1127 (1984).
26. Farrar, J. *et al. J. Immun.* **125**, 2555-2558 (1980).
27. Nagao, K., Yokoro, K. & Aaronson, S. A. *Science* **212**, 333-335 (1981).
28. Yung, Y., Eger, R., Tertian, G. & Moore, M. A. S. *J. Immun.* **121**, 794-799 (1981).
29. Bloom, G. & Kelly, J. W. *Histochemistry* **2**, 48-57 (1960).
30. Dacie, J. V. & Lewis, S. N. *Practical Hematology* 5th edn (Churchill-Livingstone, Edinburgh, 1975).
31. Fazakas de St. Groth, S. & Scheidegger, D. *J. immun. Meth.* **35**, 1-21 (1980).

Multiple sclerosis and human T-cell lymphotropic retroviruses

Hilary Koprowski*, Elaine C. DeFreitas*, Mary E. Harper†, Magnhild Sandberg-Wollheim‡, William A. Sheremata§, Marjorie Robert-Guroff†, Carl W. Saxinger†, Mark B. Feinberg†, Flossie Wong-Staal† & Robert C. Gallo†

* The Wistar Institute, 36th and Spruce Streets, Philadelphia, Pennsylvania 19104, USA

† Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

‡ Department of Neurology, University Hospital, University of Lund, S-22185 Lund, Sweden

§ Department of Neurology, University of Miami School of Medicine, PO Box 016960, Miami, Florida 33101, USA

A combination of different types of data suggests that some multiple sclerosis patients respond immunologically to, and have cerebrospinal T cells containing, a retrovirus that is related to, but distinct from, the three types of human T-cell lymphotropic viruses. The role of this virus in multiple sclerosis is uncertain.

ALTHOUGH viral infection has been postulated to play a part in multiple sclerosis (MS), a demyelinating disease of the central nervous system (CNS), there is no formal proof that any virus is involved in the aetiology of the disease. Since the discovery of the first human retroviruses¹, we have become interested in

their role as possible causes of various human diseases, including MS.

MS has some features that are characteristic of autoimmune disease. For example, it is associated with increases in absolute numbers of T cells in the cerebrospinal fluid (CSF)² and

occasionally by disturbances in the relative proportions of T-cell subpopulations during exacerbation or remission of the disease^{2,3}. A high level of oligoclonal immunoglobulin in MS CSF is a characteristic clinical finding⁴, and areas of demyelination in the brain (plaques) are usually associated with mononuclear infiltrates and often occur perivascularly⁵. Some characteristics of human retrovirus infection parallel features of autoimmune disease. The human T cell is the main target of infection of all known human retroviruses. T4-positive (helper-inducer) lymphocytes are transformed by human T-cell lymphotropic virus (HTLV)-I and -II⁶ and are killed by HTLV-III⁷, resulting in subpopulation imbalances. Infections with HTLV-I and HTLV-III have occasionally been associated with polyclonal B-cell activation and immunoglobulin production⁶. Moreover, T cells isolated from CSF of MS patients and maintained in culture closely resemble HTLV-transformed human T cells in their ability to express receptors for interleukin-2 (IL-2) for prolonged periods^{6,8}.

Recently, it was shown that HTLV-III can be detected in the brain⁹ and isolated from CSF and neural tissue of AIDS (acquired immune deficiency syndrome) patients¹⁰. Although the target cells await identification and may be of either neural or haematopoietic origin, the CNS pathology seen in MS involves cells of both lineages.

To study the possible involvement of human retroviruses in MS, we examined two sample populations. The first comes from Sweden, traditionally an area with a high incidence of MS; a recent large study reported an incidence of 5.3 per 100,000, with a female/male ratio of 1.5–1.6 and a peak age of onset of 20–39 yr¹¹. Although there is no evidence for Sweden, the uneven geographical distribution and the increasing prevalence of MS in neighbouring Finland have raised the possibility of a real increase in disease incidence¹². In western Norway, a threefold increase in prevalence from 20.1 to 59.8 per 100,000 was reported for the period 1963–83, reflecting a corresponding increase in incidence of new cases from 1.5 per 100,000 between 1953 and 1962 to 3.5 per 100,000 between 1968 and 1977 (ref. 13).

The second sample population comes from Key West, Florida, which has only recently become a high-risk area. Key West, an island 1.5×3.5 miles in size located 156 miles south-west of Miami, has a population of 26,450. Approximately 4,000 individuals are native-born. Although the first case of MS was diagnosed in Key West 57 years ago, it was not until 1981 that an unusual prevalence of the disease was suspected¹⁴.

We have examined the CSF and sera of MS patients from these two populations for the presence of antibodies reactive with antigens of the HTLV family. The availability of CSF T-cell lines from MS patients allowed us to screen for human retroviruses known to be T-lymotropic by an assay that detects viral RNA in cells in which viral genes are expressed¹⁵.

Swedish patients

We examined paired samples of CSF and serum, in each case obtained on the same day, from 35 Swedish patients with definite MS, based on the criteria of McDonald and Halliday¹⁶. These patients were between 19 and 57 years old when the samples were obtained; 14 were male and 21 were female. The disease duration ranged from 10 days to 21 years (median 8 yr) from initial onset of symptoms. In 15 patients, optic neuritis marked clinical onset of the disease. Four patients were incapacitated, 11 had some symptoms and 20 were able to carry on normal activities. Samples from 15 patients were obtained during exacerbation, which had started between 5 days and 3 months previously. Fourteen other patients were in remission and six patients had progressive disease. One patient was receiving adrenocorticotrophic hormone (ACTH) therapy at the time of sampling and two patients had received ACTH 5 and 6 months before sampling, respectively.

Paired samples of CSF and serum were obtained on multiple occasions during exacerbation, remission and progressive disease from 3 of the 35 patients, who were aged 11, 24 and 34 yr at the onset of disease. This study was limited to three patients

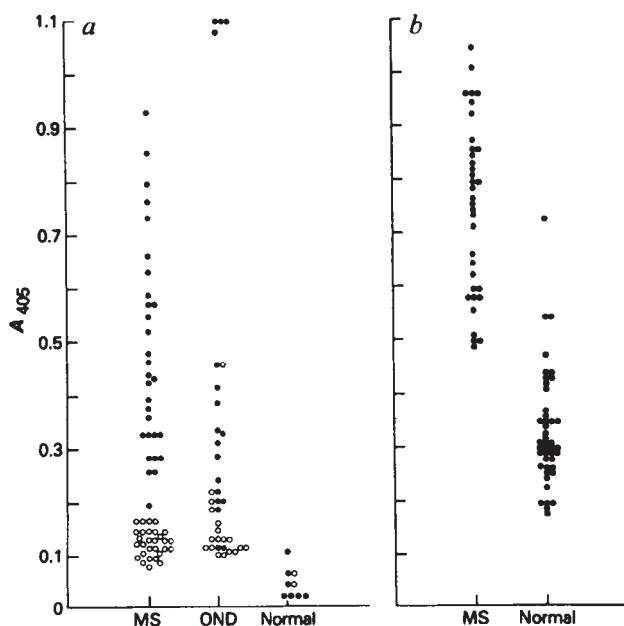


Fig. 1 *a*, Presence of antibodies reactive with HTLV-I p24 gag protein (●) and AMV p27 (○) in CSF from Swedish MS patients, patients with other neurological diseases (OND), and normal donors. All samples were tested simultaneously in the same laboratory. Data analysis by the Mann-Whitney *U*-test indicated that the MS CSF samples differed significantly from all OND CSF samples ($P=0.059$) and from OND samples taken from patients with an intact blood-brain barrier ($P=0.002$). *b*, Presence of antibodies reactive with HTLV-I p24 in sera of 35 Swedish MS patients and 43 normal Swedish donors. Serum samples from each group were diluted 1:150 in phosphate-buffered saline (PBS)/1% bovine serum albumin (BSA) and assayed for IgG antibody against HTLV-I p24 by ELISA as described for *a* below. Data analysis by the Mann-Whitney *U*-test indicated that the MS sera differed significantly from the normal sera ($P<0.001$). Results are the mean A_{405} for duplicate determinations.

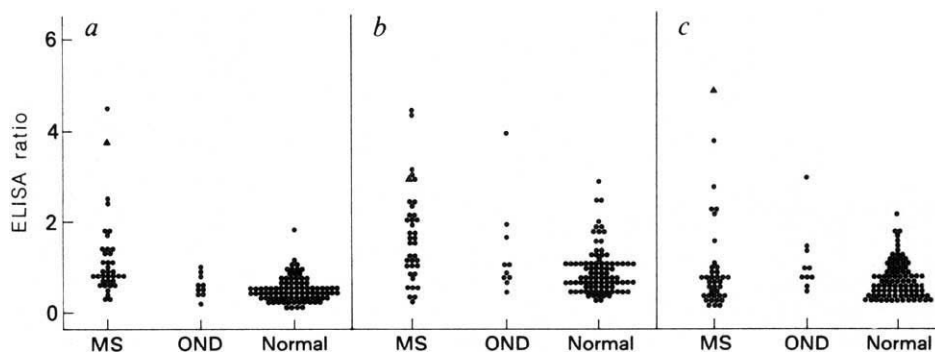
Methods. *a*, A single CSF sample from each patient was tested by ELISA using $2 \mu\text{g ml}^{-1}$ p24 or p27 bound to Immulon II plates (Dynatech, Alexandria, Virginia) in carbonate-bicarbonate buffer, pH 9.9. Nonspecific binding was blocked by addition of 1% BSA in PBS. All patient CSF samples were diluted so that the concentration of IgG equalled that found in normal CSF, then applied to the plates. After 1 h at 23 °C, an optimal concentration of peroxidase-conjugated goat anti-human IgG (heavy and light chain; KPL, Gaithersburg, Maryland) was added for 1 h at 23 °C. After extensive washing, the plates were incubated with 2,2-azino-di-(3-ethyl-benzthiazoline) sulphonate and hydrogen peroxide and the plates read at 405 nm in a Titertek Multiscan (El Fab, Helsinki, Finland). Assay conditions were standardized using HTLV-I and HTLV-III-positive sera and normal control sera. Each assay included serum samples from an HTLV-I leukaemic patient (F0920) as a positive control and NIH HTLV-II-negative sera as a negative control.

because it involves multiple taps to obtain CSF samples in sufficient quantity to perform all assays. One of the patients received ACTH several times during the sampling (see legend to Fig. 3b).

In addition, a single CSF sample was obtained from each of 18 patients with neurological diseases other than MS (OND) and from 7 healthy subjects. As a control, serum samples were also obtained from 80 healthy Swedish hospital staff members and from blood donors.

Single CSF samples from each of 35 Swedish MS patients, 18 OND patients and 7 healthy subjects were evaluated for the presence of antibody reacting with the major core protein, p24, of HTLV-I¹⁷ by enzyme-linked immunosorbent assay (ELISA), as described in Fig. 1 legend. CSF from MS patients had higher levels of the anti-p24 antibody than CSF from normal individuals and from most of the OND patients (Fig. 1a). The three OND CSF samples that were notably reactive with p24

Fig. 2 Comparison of reactivity in ELISA of HTLV-I (a), -II (b) and -III (c) disrupted virions with serum antibodies of 35 MS patients, 10 OND patients and 80 normal blood donors (all Swedish)^{17,18} (●). All samples were tested simultaneously in the same laboratory. Additional tests for specificity of all sera were performed in two ways: (1) by competition assays using extracts of virus-producing cells, phytohaemagglutinin (PHA)-stimulated normal human peripheral blood lymphocytes, and fetal calf serum¹⁹; (2) by using HTLV-III type-specific antibody and control normal antibody¹⁷. ▲, Sera were successfully competed in either assay by more than 50%.



were from a Guillain-Barre patient, a patient with meningitis of unknown origin and a patient with lymphocytic meningo-radculitis¹⁸. None of the MS CSF samples showed reactivity with the p27 of avian myeloblastosis virus (AMV) (Fig. 1a), which has low sequence homology with HTLV-I p24¹⁷, demonstrating the specificity of the antibodies for HTLV-I p24. Of the 18 OND samples tested, only the CSF from the meningo-radculitis patient was reactive with AMV p27.

ELISA analysis^{19,20} for HTLV-I p24 antibodies revealed significantly higher levels ($P < 0.005$) of these antibodies in sera from 35 MS patients (mean $A_{405} = 0.746 \pm 0.155$) than in 43 normal sera (mean $A_{405} = 0.325 \pm 0.102$) (Fig. 1b). For comparison, HTLV-I leukaemic serum F0920 routinely showed ELISA reactivity of A_{405} 1.55–1.57 in the same assay at the same serum dilution. To determine the specificity of the anti-p24 antibody in these sera, two assays were performed: one used a competitive radioimmunoassay (RIA) and the second a competitive ELISA. HTLV-I- and HTLV-III-positive sera were used to establish effective conditions to compete anti-p24 antibody with

disrupted virions of HTLV-I and of HTLV-III. In competitive-inhibition RIA and ELISA using serum samples from six MS patients, anti-p24 antibodies were completely inhibited by HTLV-I p24 and partially inhibited by disrupted HTLV-I and -III virions but not by AMV p27 (Table 1). In contrast, for serum from an HTLV-I leukaemic patient, anti-p24 antibody was completely inhibited by HTLV-I virions but not by HTLV-III virions.

The sera from the three OND patients whose CSF reacted with HTLV-I p24 also showed reactivity with HTLV-I p24. Whereas the reactivity of the case of lymphocytic meningo-radculitis¹⁸ could be considered nonspecific because the CSF reacted also with AMV p27, the case of Guillain-Barré may be linked with HTLV infection and it would be worthwhile to study more cases of the disease. The third OND case with meningitis of unknown origin could not be classified clinically.

Single serum samples from 35 MS patients, 10 patients with OND and 80 normal blood donors were evaluated by ELISA for reactivity with disrupted virions of HTLV-I, -II and -III²¹ (Fig. 2). The reactivity with HTLV-I of sera from MS patients (mean ratio 1.17) was significantly higher (Mann-Whitney U -test, $P < 0.003$) than that of OND sera (mean ratio 0.59) and of normal sera (mean ratio 0.45, $P < 0.001$). The reactivity of MS sera with HTLV-II (mean ratio 1.63) did not differ significantly from that of OND sera (mean ratio 1.24, $P < 0.08$) but did differ significantly from that of the normal sera (mean ratio 0.87, $P < 0.001$). Serum reactivity for HTLV-III showed no significant differences between all groups. Serum antibody in three MS patients was shown to be significantly inhibited by HTLV-I, -II or -III antigen or by specific antibody (see Fig. 2). No serum from OND or normal individuals was successfully competed with in these assays.

Serial CSF and serum samples from three MS patients, who were seen from the onset of their disease, were tested for antibody reactivity with HTLV-I p24 (Fig. 3). In the case of patient 01-12 (Fig. 3a), both the CSF and serum anti-p24 antibody levels fluctuated markedly between 1982 and 1984. A peak of CSF antibody in February 1983 was accompanied by elevation of the serum antibody, but the peaks of serum antibody observed in August 1983 and December 1984 were not accompanied by a rise in CSF antibodies. All antibody detected was of the IgG class; no IgM antibody was detected. The total IgG concentration in serum and CSF remained unchanged during the observation period (Fig. 3a); thus, the increase in specific antibody could not be explained by polyclonal increases in IgG levels.

In the second MS patient, 20-01 (Fig. 3b), p24 reactivity was detected in serum and CSF samples collected from 1977 to 1982; the two curves are parallel except from November 1981 onwards, when a rise in CSF antibody was accompanied by a decrease in serum antibody. Again, all antibody was IgG; no IgM was detected. The fluctuation in anti-p24 antibody concentration did not reflect changes in CSF and serum IgG levels, which remained relatively constant.

Table 1 Competitive inhibition for HTLV-I p24 antibodies in sera of Swedish MS patients

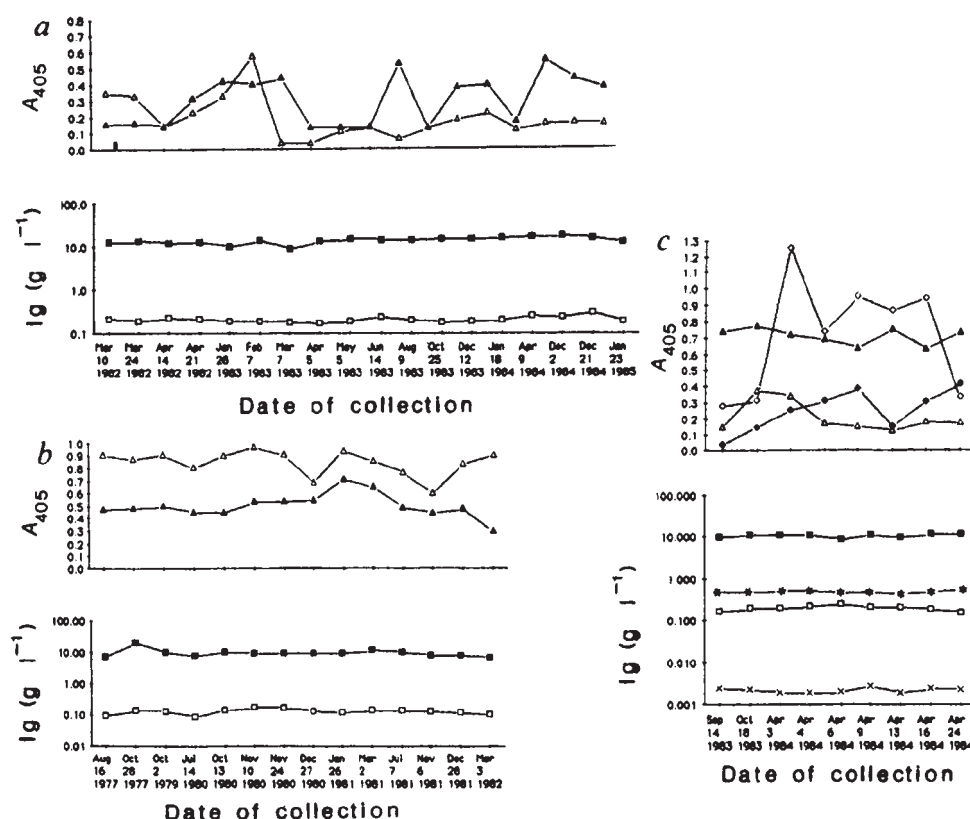
Competing antigen: Patient	% Inhibition of HTLV-I p24 antibody			
	Disrupted virions HTLV-I	HTLV-III	HTLV-I p24	AMV p27
13-10*	46	44	86	6
01-12*	29	30	83	4
3380†	58	51	98	NT
2268†	40	31	83	-8
2167†	61	29	89	NT
1738†	29	15	86	NT
F0920‡	90	28	94	NT

* Patients' sera were diluted 1:25 and incubated overnight on plates precoated with goat anti-human IgG as a 'catcher' antibody. Inhibition of HTLV-I p24 antibody was determined by reacting various concentrations of disrupted virions of HTLV-I and -III, AMV p27 and HTLV-I p24 with wells containing patients' sera for 3 h at 37 °C. After washing, binding of HTLV-I p24 was assessed by incubating all cells with 0.25 μ g ¹²⁵I-labelled HTLV-I p24 (1×10^5 c.p.m. per well). Inhibition is expressed as $[(^{125}\text{I-p24 c.p.m. bound with no competing antigen} - ^{125}\text{I-p24 c.p.m. bound with competing antigen}) / \text{total } ^{125}\text{I-p24 c.p.m.}]^{-1} \times 100$. Values are the percentage inhibition of ¹²⁵I-p24 binding by 0.5 μ g of competing antigen.

† Specificity of antibodies against HTLV-I p24 in four MS sera was determined by competitive inhibition in an ELISA with soluble HTLV-I p24, disrupted virions of HTLV-I and -III and AMV p27. Equal volumes of patients' sera (1:100 dilution) were incubated with soluble antigen (0.7 μ g ml⁻¹) for 30 min then applied to plates precoated with HTLV-I p24 as described in Fig. 1 legend. Per cent inhibition was calculated as follows: $[(A_{405} \text{ of sera without soluble competitor} - A_{405} \text{ of sera with competitor}) / A_{405} \text{ competition}] \times 100$. NT, not tested.

‡ Serum from an HTLV-I leukaemic patient was diluted 1:500 and inhibitors used at 0.35 μ g ml⁻¹.

Fig. 3 Antibody reactivity with HTLV-I p24 in serial samples of CSF and sera obtained from Swedish MS patients during their disease, compared with concentrations of IgG and IgM in the same CSF and sera. All samples were tested simultaneously in the same laboratory. **a**, Patient 01-12: onset of MS in September 1977 at age 34 yr. The course of the disease was characterized by frequent, acute and short exacerbations, particularly during the study period from 1982 to 1984. The patient received no ACTH or steroid treatment. **b**, Patient 20-01: onset of MS in November 1976 at age 24 yr. The course of the disease was characterized by exacerbations from its onset to 1983, with slow progression of the disease thereafter. The patient received two ACTH treatments in December 1980 and December 1981. **c**, Patient 13-10: onset of MS in March 1976 at age 11 yr with exacerbations in 1979, 1980, April and September 1983, and April 1984, with complete recovery between exacerbations. Patient treated with ACTH only during April 1983, 1 yr before the exacerbation in April 1984. Δ , IgG antibody to HTLV-I p24 in CSF; $\blacktriangle$, IgG antibody to HTLV-I p24 in sera; $\square$, $\blacksquare$: concentration of IgG in CSF and sera, respectively; $\times$, $*$: concentration of IgM in CSF and sera, respectively.



Methods. The assay for HTLV-I p24 antibodies was as described in Fig. 1 legend. The immunoglobulin concentrations were determined by an electroimmunoassay described elsewhere²². Serum samples were diluted for the ELISA assay to achieve approximately the same immunoglobulin concentration as found in the CSF samples in order to compare the relative amounts of specific antibody. The immunoglobulin concentrations plotted are therefore those used in the assay.

In the third patient sampled, 13-10 (Fig. 3c), it was possible to follow the anti-p24 antibody levels over several months before and during an acute clinical exacerbation in April 1984. A dramatic rise in CSF anti-p24 antibody of the IgM class was detected between 14 September 1983 and 3 April 1984, accompanied by a lesser but still significant rise in CSF IgG antibody to p24. During April 1984, the CSF IgG antibody decreased to its original value. The CSF IgM antibody levels remained high until 13 April 1984, after which they returned to the levels observed before exacerbation. The concentration of IgM serum antibody fluctuated, whereas that of the IgG serum antibody remained relatively high during the observation period. None of these changes can be explained by concomitant changes in CSF or serum immunoglobulin concentrations, which remained relatively stable during the observation period.

Intrathecal production of immunoglobulin

IgG and albumin concentrations in CSF and serum from the Swedish MS patients were determined²² to obtain the IgG-albumin index (CSF IgG $\times$ serum albumin / serum IgG $\times$ CSF albumin), an indicator of intrathecal immunoglobulin production in CSF²³. Values above 0.66 indicate intrathecal immunoglobulin production in patients with an intact blood-brain barrier²³. For the three patients who were repeatedly sampled, the indices ranged from 0.81 to 6.82, with only two samples showing values <1.00 . For the MS patients in Fig. 1, values ranged from 0.50 to 2.16, with only four values less than 0.66. Thus, a portion of the CSF IgG of these patients is synthesized in the CNS and does not originate in the blood.

The ratio of CSF albumin to serum albumin is taken as an indicator of the integrity of the blood-brain barrier²⁴; this ratio

was within the normal range for all MS patients, indicating an intact blood-brain barrier. However, the three OND patients with high levels of CSF antibody reactive with HTLV-I p24 (Fig. 1a) had ratios of 0.033, 0.020 and 0.011, compared with 0.0074 for normal donors, indicating that the blood-brain barrier was not intact, permitting leakage of serum proteins into the CSF.

Patients from Key West, Florida

At the time of assessment on 1 September 1983, 37 patients with MS were identified in Key West¹⁴. Only one was not Caucasian, and there were two pairs of siblings with MS and two patients with MS-afflicted siblings living elsewhere. The diagnosis of MS was again based on the criteria of McDonald and Halliday¹⁶. Of the 37 patients, 17 were available to us for this study. Five were born in Key West; of the other 12, 5 had moved there at ages of 5–18 yr and 7 at age 22–39 yr. The onset of MS preceded this study by 1–5 yr in 12 patients, by 9 yr in one patient, and by 14, 16, 19 and 20 yr in the remaining patients. Three patients had never received steroid therapy, 13 had received steroid or ACTH therapy (no later than 1983) and one patient was being treated with Immuran at the time of study.

Paired samples of CSF and serum were available from 9 of the 17 patients, and serum samples were available from the remaining 8 patients. A second serum sample obtained 4 weeks later was available from 9 of the 17 patients. Seventeen close contacts of the patients (spouse, parent, sibling) and 17 other individuals also donated blood for this study.

Table 2 shows the number of sera from MS patients and from their contacts that were reactive with HTLV-I p24 and disrupted virions of HTLV-I, -II and -III at one point in time. Seven serum

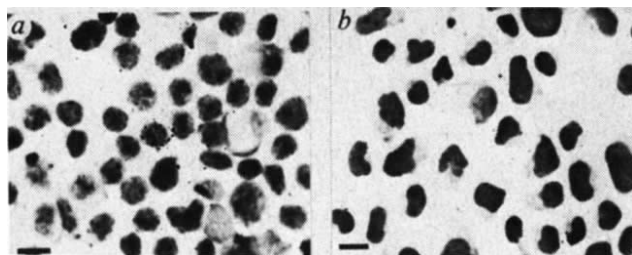


Fig. 4 Detection of cells expressing RNA with homology to HTLV-I in cultures derived from CSF of MS patients. *a*, Representative field of cultured cells obtained originally from CSF of native-born Key West MS patient no. 4 and hybridized with an HTLV-I 3' probe. This patient showed onset of MS in 1982 at age 21 yr. The disease followed a relapsing-remitting course with accumulated moderate neurological deficits attributed to spinal cord involvement; there was major manifestation of mild to moderate ataxia and spastic paraparesis. Labelled cells were present at a frequency of 0.01%. *b*, Representative field of cultured cells obtained from CSF of a healthy Swedish subject and hybridized with the HTLV-I 3' probe, as described for *a*. No labelled cells present. Pictures of representative fields of cultured cells obtained from two other Key West MS patients and one Swedish patient (see text) were similar to *a*. Scale bar = 10 μ m.

Methods. Freshly isolated CSF cells were co-cultured in cluster plates (Costar) with autologous γ -irradiated peripheral blood mononuclear cells in complete medium containing purified PHA (PHA-P, 1 μ g ml⁻¹; Burroughs-Wellcome) for 3 days. Cell concentrations were adjusted to 2×10^5 ml⁻¹ in complete growth media with 50 ng ml⁻¹ recombinant IL-2 (Sandoz, Vienna, Austria). Cells were maintained in logarithmic growth by periodic supplementation with fresh media containing IL-2 and by stimulation with PHA-P. All cell lines were maintained in the same laboratory before hybridization. *In situ* hybridization was carried out using ³⁵S-labelled RNA probe specific for the 3' region of HTLV-I¹⁵. The RNA probe was transcribed from a full-length genomic clone of HTLV-I and averaged 1–2 kilobases long. Probe was hybridized at $1\text{--}2 \times 10^8$ d.p.m. ml⁻¹ in relaxed conditions (45 °C) and autoradiographed for 4–8 days. All cell lines were hybridized in the same conditions in the same laboratory, and cells were examined using a double-blind code.

samples from MS patients were reactive with HTLV-I, four with HTLV-II and one with HTLV-III; three sera also reacted with HTLV-I p24. Serum of one patient reacted with all three HTLV types, whereas serum of another patient reacted with only HTLV-I and -II. Of the 17 MS sera, 10 were reactive with an HTLV antigen at least once. Of the 17 donors who had intimate contact with the MS patients, two showed serum reactivity with HTLV-III: one of the sera, obtained from a homosexual brother of one of the MS patients, was HTLV-III-seropositive; his serum reactivity might reflect true exposure to the HTLV-III agent of AIDS rather than a cross-reactivity. The second donor was a male not known to be in any HTLV high-risk group. None of the sera from 17 people employed at the hospital where most of the MS cases were treated showed reactivity with HTLV-I, -II or -III.

Of nine CSF samples available from these MS patients, only one sample was reactive with HTLV antigens, specifically with HTLV-I p24. Of the eight negative samples, seven were from patients whose sera were reactive with HTLV antigens.

HTLV-I-related nucleotide sequences

In situ hybridization was used to examine cultured CSF cells obtained from five Key West MS patients, three Swedish MS patients and two healthy Swedish subjects for the presence of HTLV-specific RNA by methods described previously (ref. 15 and unpublished results with M. Feinberg), using specific molecularly cloned HTLV-I and HTLV-III probes. So far, three of the CSF cultures derived from Key West MS patients (who had suffered from MS for 3, 4 and 21 yr) and one CSF culture derived from a Swedish MS case (who had suffered from the disease for 1 month) have shown the presence of labelled cells

following hybridization under low-stringency conditions (45 °C instead of 50 °C) with the HTLV-I 3' RNA probe, which contains predominantly long terminal repeat sequences. These positive cells exhibited an average of 50 grains per cell (Fig. 4*a*) and were observed at a frequency of from 0.001 to 0.01% of the 5×10^5 cells generally examined. In contrast, hybridization of the same HTLV-I probe under identical conditions to cultured CSF cells from two normal individuals consistently failed to reveal labelled cells (Fig. 4*b*); similarly, there were negative results with normal peripheral blood cells, bone marrow and other cell types. Hybridization of an HTLV-I probe specific for the *gag* coding region to CSF cells from two Key West MS patients also revealed rare labelled cells. Thus, cells in four MS CSF cultures contain RNA that hybridizes with HTLV-I under low-stringency conditions. In several experiments in which hybridization was carried out under standard stringency conditions (50 °C), no labelled cells were observed in CSF from one Key West MS patient. Although difficult to quantitate because of the low number of positive cells observed after hybridization at 45 °C, the lack of detectable label under conditions of higher stringency suggests that the RNA detected shares only partial homology with the HTLV-I genome. No hybridization was detected in the same MS CSF cultures using an HTLV-III 3' RNA probe under reduced stringency, indicating a lack of RNA homologous to this virus in the cultured cells.

Discussion

The data presented here indicate the presence of antibodies reactive with HTLV antigens in CSF of Swedish MS patients at concentrations significantly higher than those in CSF samples from Swedish OND patients and from healthy donors. The three OND patients with high CSF levels of anti-HTLV antibodies were shown to have an impaired blood-brain barrier, which suggests that the antibodies were derived from serum. The CSF from seven healthy subjects with intact blood-brain barriers showed no reactivity. Antibodies reactive with HTLV-I p24 antigen (Fig. 1*b*) and with disrupted virions of HTLV-I and -II (Fig. 2) were at higher levels in sera of Swedish MS patients than in sera from OND patients (Fig. 2) or from normal healthy subjects (Figs 1*b*, 2). Perhaps even more significant is the fact that of 17 sera collected from Key West MS patients at various times, 10 reacted at one time or another with an HTLV antigen, whereas reactivity was observed in only 2 of 17 sera obtained from close contacts of the patients (one of whom was a homosexual with serum antibody against HTLV-III) and in none of 17 sera from hospital staff workers from the same area.

One cannot exclude the possibility that, as in the case of measles antibodies in CSF of MS patients²⁵, the HTLV antibodies in CSF represent only a small fraction of the immunoglobulin produced intrathecally by MS patients. Unlike measles antibodies²⁶, however, the level of HTLV antibodies fluctuated in the CSF of MS patients from whom serial CSF samples were available. In one case studied (Fig. 3*c*), a sharp rise of these antibodies accompanied an exacerbation of the clinical disease. It is important, therefore, to emphasize that the antibody reactivity identified in some MS sera or CSF was detected at some periods but not at others.

No uniform response to one particular type of HTLV was observed; this was especially obvious in the Key West MS patients, in whom serum reactivity with any of the three types of HTLV varied among patients. In two cases, the same serum samples reacted with more than one type of HTLV (Table 2). The HTLV-I anti-p24 reactivity in six Swedish MS sera was comparably inhibited by HTLV-I and HTLV-III disrupted virions (Table 1), which may reflect recognition of an epitope(s) on HTLV-I p24 which is shared with other types of HTLV²⁷. The fact that none of the Swedish MS CSF samples reacted with AMV p27, and that 10 HTLV-reactive sera from Key West patients did not react with disrupted virions of visna virus or gibbon ape leukaemia virus indicates that the antibody response of the MS patients is directed specifically towards one or more HTLV-related proteins. But as our data suggest that it is not

Table 2 HTLV-reactivity of sera from Key West patients, intimate contacts and hospital workers

Clinical status	No. of patients	No. of sera reacting with:				No. of positive cases/total*
		HTLV-I p24	HTLV-I	Disrupted virions of HTLV-II	HTLV-III	
MS	17	3	7	4	1	10†/17
Contacts	17	1	0	0	1	2‡/17
Hospital workers	17	0	0	0	0	0/17

Reactivity was determined by ELISA; the results were all confirmed by competition with either antigens of or specific antibodies to HTLV-I, -II and -III¹⁷⁻¹⁹. All samples were tested simultaneously in the same laboratory. The contacts and hospital workers groups were each compared with MS patients using Yates' χ^2 analysis: $\chi^2 = 5.44$ ($P = 0.020$) for contact group; $\chi^2 = 9.67$ ($P = 0.002$) for hospital workers.

* The positive results in the MS patients are transient; they are found at only a specific point in time in any given patient.

† Serum of one patient reacted with all three types of HTLV; serum of another patient reacted with types I and II, and serum of one patient reacted only with HTLV-I p24.

‡ One case was a homosexual brother of an MS patient.

directed exclusively at any of the three known HTLV types, it is difficult to apply the rigid criteria established for evaluation of serological results in diseases linked unquestionably with either HTLV-I (leukaemia, lymphoma) or HTLV-III (AIDS). Indeed, interpretation of serological results in the present study is more arbitrary and the detection of HTLV antibodies should in no way be considered of diagnostic value for MS.

The detection of RNA homologous to HTLV-I in CSF cells of four MS cases whose sera reacted with HTLV-I antigen (Table 2) indicates that a human retroviral genome is being expressed and suggests the possibility that a virus is present. The low frequency of hybridizing cells is similar to the situation observed in primary samples of T cells from HTLV-III-infected patients (AIDS and AIDS-related complex, ARC), where 0.01% or less of cells express RNA¹⁵. The *in vitro* cultivation of T cells infected with HTLV-III results in increased expression of RNA and protein; the actual percentage of cells expressing virus after *in vitro* cultivation is probably related to the frequency of virus-infected cells in the original sample. *In situ* hybridization studies with another retrovirus, visna virus, indicated that fewer than 1% of cells expressing low amounts of viral RNA were found in the tissue samples taken from infected sheep.^{28,29}

As the presence of anti-HTLV antibodies in the CSF of MS patients suggests an antigenic stimulation from the CNS, it seems reasonable to search in the CNS for the presence of the appropriate antigen. The only evidence that the human brain is susceptible to HTLV infection is the detection of HTLV-III in brains of children and adults with AIDS encephalopathy¹⁰. In diseases caused by HTLV, the target for primary virus infection is the peripheral T cell. If the same is true in MS, infected peripheral T cells might transport the virus to the CNS, cells of which might be permissive for viral replication. Damage to myelin-producing glial cells could result directly from viral infection. Whereas no evidence exists for the presence of virions in MS brain, it should be stressed that a specific search for the presence of human retrovirus or its components in brain cells has not yet been made.

The presence of CSF cells that hybridize with the HTLV-I probe in MS patients but not in healthy subjects may suggest that T cells present in CSF become infected during the disease. This hypothesis is strengthened by the fact that some T cells obtained from CSF of MS patients share characteristics with HTLV-transformed T cells. For example, like HTLV-transformed T cells⁶, one clone from CSF cells of a Swedish MS patient had a low, but significant, number of cells with lobular or multiple nuclei⁵. In addition, most mitogen-induced MS CSF T-cell lines consisted predominantly of OKT3,4-positive cells⁸, as did cells isolated from HTLV-associated T-cell malignancies and cord blood lymphocytes transformed *in vitro* with HTLV-I⁶. Finally, most MS CSF T-cell lines grow to saturation densities of $\geq 2 \times 10^6$ ml⁻¹ (ref. 8), like the HTLV-transformed lymphocytes⁶, and grow for extended periods of time in exogenous IL-2 only⁸. The exception is the one clone mentioned above, which is IL-2 receptor-positive but neither produces nor depends

on IL-2 for growth; in this regard, these cells resemble both cord blood and bone marrow-derived lymphocytes transformed by HTLV-I^{30,31}.

If an HTLV-related retrovirus is at all involved in the aetiology of MS, how does one account for three of our findings: (1) the serological detection of anti-HTLV antibodies in less than 70% of the patients; (2) the consistently low level of serum antibody responses in those sera showing some reactivity; and (3) some variation in specificity of the antibody response to HTLV-I, -II or -III? Assuming MS is a disease caused by one agent, the first observation might be explained by the fluctuation in antibody levels during the disease, so that antibodies might not be detected in a single CSF or serum specimen. Alternatively, all MS patients might be infected but only a proportion might respond with the production of antibodies that cross-react with HTLV, or the agent might consistently fail to induce detectable amounts of antibodies. In fact, HTLV-III infection does not induce an antibody response in all infected subjects although the proportion is around 4% (refs 32, 33). Regarding point (2), the low titre of antibodies could be due to transient and/or low-level expression of the putative virus, or to its cross-reactivity with the known HTLV types being very slight. Finally, MS may consist of several subtypes with different aetiologies and the infectious HTLV-related agent may be a factor in only some subtypes. Our data indicate that the agent detected in this study is not HTLV-I, -II or -III. A related but distinct retrovirus might be expected to cross-react weakly with one or more antigenic and genomic determinants of the known HTLVs, just as a weak cross-reactivity with HTLV-I probes led to the discovery of HTLV-II^{34,35}.

The variation in the reactivity of the antibodies with HTLV-I, -II or -III may be explained if the epitopes recognized by these antibodies partially cross-react with the known HTLV types. A novel HTLV related to HTLV-I would be expected to react with HTLV-II antigens in view of the substantial homology between these two retroviruses^{34,35}. Furthermore, some antibodies might be more reactive with HTLV-II proteins than with HTLV-I if the new HTLV had greater homology with HTLV-II than with HTLV-I. However, homology cannot be used to explain the two serum samples from MS patients which were reactive with both HTLV-III and HTLV-I or HTLV-II (Fig. 2, Table 2), as the genome of HTLV-III differs substantially from the other two types. This could conceivably be explained by an HTLV with the properties of a recombinant of the two major types, with variation in expression of specific viral antigen(s) among different patients and even in the course of the disease in a given patient.

The incubation period for HTLV-I and -II is 5–30 yr^{34,36}. A similar incubation period for the putative HTLV-related retrovirus would be consistent with the notion of exposure in early adolescence to any infectious agent that might be involved in the aetiology of MS³⁷.

We have discussed our data as though they indicate the presence and involvement of a novel HTLV-related virus in MS.

If that is the case, our attempts to isolate the virus should eventually be successful. For now, the data are suggestive, but not conclusive.

This work was supported by NIH grants NS-11036 and AI-19987-01, and by grants from the Reynolds Foundation and the Swedish MRC (project no. B84-19X-06265-03B). We thank

Received 4 July; accepted 2 October 1985.

1. Poiesz, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7415-7419 (1980).
2. Sandberg-Wollheim, M. *Scand J. Immun.* **17**, 575-581 (1983).
3. Cashman, N., Martin, C., Eizenbaum, J. F., Degos, J. D. & Bach, M. A. *J. clin. Invest.* **70**, 387-392 (1982).
4. Link, H. *et al. Acta neurol. scand.* **55**, 173-189 (1977).
5. Guseo, A. & Jellinger, K. J. *Neurol.* **211**, 51-60 (1974).
6. Gallo, C. *Cancer Surv.* **3**(1), 113-159 (1984).
7. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
8. DeFreitas, E., Sandberg-Wollheim, M., Schonely, K. & Koprowski, H. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
9. Shaw, G. M. *et al. Science* **227**, 177-182 (1985).
10. Ho, D. D. *et al. New Engl. J. Med.* (in the press).
11. Broman, T., Andersen, O. & Bergmann, L. *Acta neurol. scand.* **63**(1), 6-33 (1981).
12. Kinnunen, E. *Neurology* **34**, 457-461 (1984).
13. Larsen, J. P., Riise, T., Kvaale, G., Nyland, H. & Aarli, J. A. *Acta neurol. scand.* **69**, Suppl. 98, 372-373 (1984).
14. Shermata, W. A., Poskanzer, D. C., Withum, D. G., MacLeod, C. L. & Whiteside, M. E. *Lancet* **ii**, 618 (1985).
15. Harper, M., Marselle, L., Wong-Staal, F. & Gallo, R. C. *Proc. natn. Acad. Sci.* (in the press).
16. McDonald, W. E. & Halliday, A. M. *Br. med. Bull.* **33**, 4-9 (1977).
17. Kalyanaraman, V. S., Sarngadharan, M. G., Poiesz, B., Ruscetti, F. W. & Gallo, R. C. *J. Virol.* **81**, 906-915 (1981).

M. G. Sarngadharan for discussions and for providing purified antigens of HTLV-I, -II and -III. Technical assistance was provided by K. Schonely, M. Boufal, A. Karaminides, L. Marselle, K. Chayt, A. Jennings, F. Laurent and Mei-Wah Hoh. Statistical analysis was performed by Mitchell Gail of the National Institutes of Health.

18. Ryberg, B., Nilsson, B., Burgdorfer, W. & Barbour, A. G. *Lancet* **ii**, 519 (1983).
19. Saxinger, W. C. & Gallo, R. C. *Lab. Invest.* **49**, 371-377 (1983).
20. Sarngadharan, M. G., Popovic, M., Bruch, L., Schupbach, J. & Gallo, R. C. *Science* **224**, 506-508 (1984).
21. Robert-Guroff, M. *et al. Science* **215**, 975-978 (1982).
22. Laurell, C.-B. *Scand. J. clin. Lab. Invest.* **29**, Suppl. 124, 21-37 (1972).
23. Olsson, J.-E. & Pettersson, B. *Acta neurol. scand.* **53**, 308-322 (1976).
24. Reiber, H. *J. Neurol.* **224**, 89-99 (1980).
25. Norrby, E. *Prog. med. Virol.* **24**, 1-39 (1978).
26. Arnadottir, T. *et al. Archs Neurol.* **36**, 261-265 (1979).
27. Sarngadharan, M. G., Bruch, L., Popovic, M. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3481-3484 (1985).
28. Bratic, M., Stowring, L., Ventura, P. & Haase, A. T. *Nature* **292**, 240-242 (1981).
29. Stowring, L. *et al. Virology* **141**, 311-318 (1985).
30. Miyoshi, I. *et al. Nature* **294**, 770-771 (1981).
31. Markham, P. D. *et al. Int. J. Cancer* **31**, 413-420 (1981).
32. Salahuddin, S. Z. *et al. Lancet* **ii**, 1418-1420 (1984).
33. Groopman, J. E. *et al. Ann. intern. Med.* **102**, 63-66 (1985).
34. Kalyanaraman, V. S. *et al. Science* **218**, 571-573 (1982).
35. Gelmann, E. P., Franchini, G., Manzari, V., Wong-Staal, F. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 993-997 (1984).
36. Takatsuki, K. *et al. Gann* **28**, 13-22 (1982).
37. Dean, G. & Kurtzke, J. F. *Br. med. J.* **3**, 725-729 (1971).

LETTERS TO NATURE

A blue stellar population in the H I bridge between the two Magellanic Clouds

M. J. Irwin*, W. E. Kunkel† & S. Demers‡

* Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

† Mount Wilson and Las Campanas Observatories, La Serena, Chile

‡ Université de Montréal, CP168, Succursale A, Montréal, Québec, Canada H3C 3J7

Radio observations have shown¹ that a common H I cloud surrounds the Magellanic Cloud system. However, no stellar link between the two galaxies has previously been found. The furthest known stellar extension eastward of the Small Magellanic Cloud (SMC) is the blue association of stars found by Kunkel², which lie close to the extreme tip of the 'wing' of the SMC. The SMC wing was discovered by Shapley³, who described it as "a large cloud of faint stars extending eastward from the SMC toward the LMC [Large Magellanic Cloud]". More recently, Westerlund⁴ described the brighter ($V < 20.0$) stellar population of the wing as being a mixture of extreme Population I, OB associations, H II regions and both blue and yellow supergiants. The rest of the Magellanic system shows a strong correlation between H I density and the presence of stellar condensations, and this has led to speculation that stellar associations may be present in the H I extension of the optical wing (the H I bridge indicated in Fig. 1) and also independently in regions of high H I density in the Magellanic stream⁵. Deep H α photography⁶ shows weak diffuse emission between the Magellanic Clouds, supporting the possible existence of an optical counterpart to the H I bridge, but star counts near the eastern tip of the wing⁷ failed to reveal any link and indeed corroborated earlier work on the structure of the wing⁸. We now report the discovery of several hundred blue main-sequence stars in an area reaching 9° east of the SMC, coinciding with the H I ridge extending from the wing. These stars lie midway between the two Clouds and are more than 2 kpc further east than what was believed to be the eastern tip of the SMC wing. The apparent magnitudes of the brightest main-sequence stars suggest an age of ~100 Myr, similar to ages of associations in the wing.

The discovery² of stellar aggregates at the extreme eastern tip of the wing prompted us to investigate a large area to the east of the wing. For this purpose we obtained a pair of UK Schmidt plates (IIIaJ + GG385 and IIIaF + OG590) centred on RA 02 h 37 min, dec. -73° 43'. The plates were scanned and processed by the Automatic Plate Measuring (APM) facility at Cambridge and calibrated from a photoelectric sequence (W.E.K., M.J.I. and S.D., in preparation). Magnitudes (R) and colours ($J-R$) were obtained for over 120,000 stellar objects in a 6° × 6° area. This number corresponds to a magnitude limit of $m_J \sim 21$ and $m_R \sim 20$.

The area covered by the plates and its relationship to the Magellanic system are shown in Fig. 1. Three levels of integrated H I column densities are shown for the LMC⁴ and the SMC⁹, corresponding to levels of 4, 8 and $\geq 12 \times 10^{20}$ H I atoms cm⁻². The optical centres¹⁰ of both galaxies are represented by dots.

The colour-magnitude diagram of the 120,000 stars reproduced in Fig. 2 reveals a well-defined blue main sequence, extending upward to $m_R \sim 15$, and well separated from the expected foreground galactic stars. Multicolour photometry¹¹

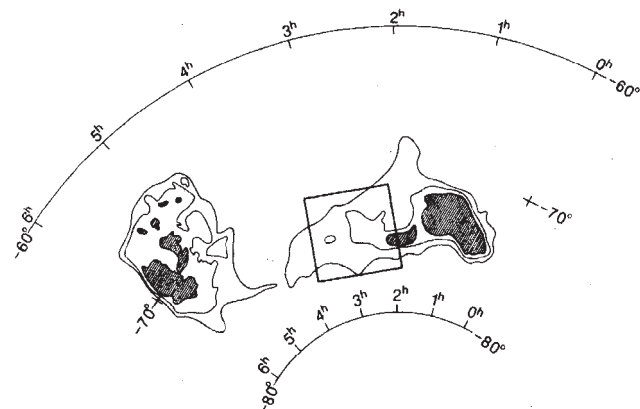


Fig. 1 The field investigated (square) relative to the Magellanic Clouds represented here by three levels of H I column density: 4, 8 and $\geq 12 \times 10^{20}$ atoms cm⁻². The dots represent the optical centres of the two galaxies.

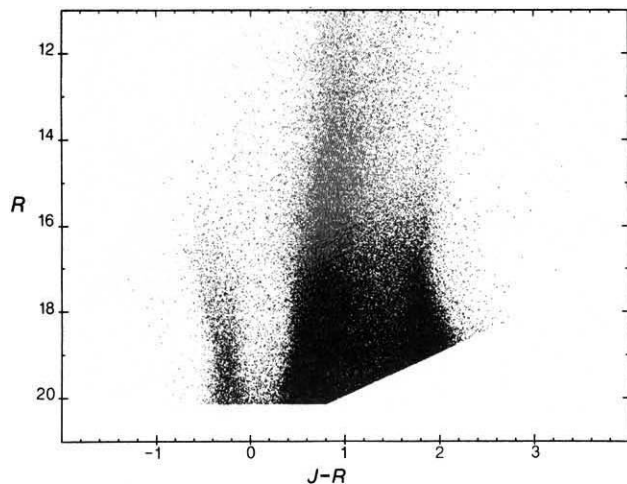


Fig. 2 A stellar colour-magnitude diagram in photographic J (B_r) and R of the $6^\circ \times 6^\circ$ of sky investigated. A Zero Age Main Sequence (ZAMS) is clearly discernible (dashed line) as are the foreground galactic halo (yellow $J-R \sim 0.8$) and disk (red $J-R \sim 1.9$) stars. Some 4,000 stars are in the region blueward of $J-R=0$.

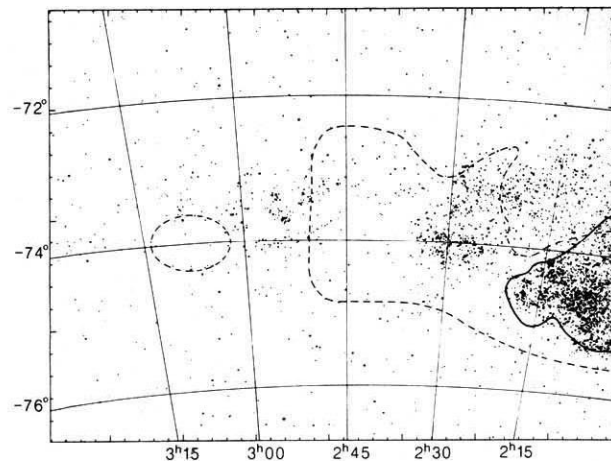


Fig. 3 The spatial distribution of the stars with colour $J-R < 0$. The tip of the wing (see ref. 8) is outlined (solid line) together with the H I isophote corresponding to 8×10^{20} atoms cm^{-2} (dashed line).

of Cepheids in the system has shown that the colour excess toward the tip of the wing is small, $E(B-V) \leq 0.05$, indicating that these blue stars have magnitudes and colours consistent with a sequence located at roughly the same distance as the Magellanic Clouds. Figure 3 gives the spatial distribution of the $\sim 4,000$ main-sequence stars with $J-R < 0$. These stars represent around 3% of the total stellar population of the 36 deg^2 area. The tip⁸ of the wing and the H I isophotes are outlined. As expected, most of the blue stars, occur in the extreme tip of the wing, seen on the right (Fig. 3). The concentration at RA 02 h 30 min coincides with the previously known stellar aggregate². We see here for the first time a large area of blue stars lying to the north of the extreme tip, and the group of blue stars at RA 02 h 52 min of angular extent $\sim 1^\circ$. This latter group is 9° from the optical centre of the SMC and 12.5° from the LMC, corresponding to distances of 10.6 and 12.5 kpc, respectively, which are well outside the radius¹⁰ of both clouds. Furthermore, Fig. 3 clearly reveals an overall optical counterpart to the H I bridge, although some of the blue stars are apparently in regions of low H I density.

Colour-magnitude diagrams for all the areas possessing a marked excess of blue stars have been analysed separately. Apart from number density of images, all the regions have almost identical blue main sequences (within ± 0.1), confirming that these blue stars indeed belong to the Magellanic Clouds. By statistically removing an equivalent foreground comparison colour-magnitude diagram, it is also possible to search the region with colours $J-R > 0$ for red giants and horizontal branch stars. Preliminary results indicate that the combined number of Magellanic red giants and horizontal branch stars is less than 10% of the number of blue main-sequence stars in the selected areas. Furthermore, for stars with $J-R > 0$, no clumpiness or overall stellar density gradient is discernible across the field, as observed⁷ for the halo region of the SMC. Even limiting the search to those stars having the magnitude and colour of the horizontal branch stars produces no definitive non-uniformity of images. This is not too surprising as these stars form less than 1% of the total population at these magnitudes and colours and are therefore swamped by the foreground component.

As a further check, number density maps of the whole $6^\circ \times 6^\circ$ area were constructed. These reveal a distinct excess of images only in the regions where the extra blue stars are found, confirming that it is the blue main-sequence stars that dominate the bridge population. Also, the noise in the number density map indicates that although the extra stars are visible with hindsight on deep photographs, two-colour photometry is

needed to identify them unambiguously. All this evidence strongly suggests that we are looking outside what is known as the SMC. There is no doubt from the individual colour-magnitude diagrams that the main-sequence stars are at a similar distance to the Magellanic Clouds. Photometry of Cepheids¹² throughout the SMC strongly suggests that the wing is pointing towards us, thus becoming progressively closer as it goes further east. The nearest well-studied¹³ cluster, Lindsay 113, lies in the wing 4.5° east of the optical centre of the SMC at a distance of 58 kpc. This is midway between the accepted distance¹² for the SMC and LMC of 62 and 54 kpc, respectively. We will adopt a preliminary distance of 58 kpc for the Magellanic bridge region, keeping in mind that recent observations¹⁴ have suggested that the distances of both Magellanic Clouds may require a major revision.

At this distance the brightest main-sequence stars are at $V \sim 15$ or $M_V \sim -4$. Adopting an age calibration¹⁵ of the Magellanic Cloud clusters leads to an age estimate of less than 100 Myr for the new blue stellar associations. This is comparable with the ages¹⁶ of clusters and associations in the wing. A study leading to a more accurate knowledge of the chemical composition, age and distance (and hence origin) of the stars in the bridge region is now under way (W.E.K., M.J.I. and S.D., in preparation). Preliminary results from the photographic study indicate that the blue stars in the bridge region have a common origin with similar blue stars in the SMC wing and that they are located midway between the two Clouds.

This study was funded in part by the SERC, UK, and Natural Sciences and Engineering Research Council, Canada. We thank the UKSTU for taking the plates and the members of the APM group for providing the services of the scanning facility. S.D. is on leave from the Université de Montréal.

Received 26 July; accepted 17 September 1985.

- Mathewson, D. S. & Ford, V. L. *IAU Symp.* No. 108, 125-136 (1984).
- Kunkel, W. E. *IAU Symp.* No. 85, 353-355 (1980).
- Shapley, H. *Harvard Obs. Bull.* **814**, 8-9 (1940).
- Westerlund, B. E. *Vistas Astr.* **12**, 335-356 (1971).
- Bruck, M. T. & Hawkins, M. R. S. *IAU Symp.* No. 108, 103-104 (1984).
- Johnson, P. G., Meaburn, J. & Osman, A. M. I. *Mon. Not. R. astr. Soc.* **198**, 985-989 (1982).
- Bruck, M. T. *Astr. Astrophys.* **68**, 181-187 (1978).
- Westerlund, B. E. & Glaspey, J. *Astr. Astrophys.* **10**, 1-7 (1971).
- McGee, R. X. & Newton, L. M. *Proc. Astr. Soc. Austr.* **4**, 189-195 (1981).
- de Vaucouleurs, G. & Freeman, K. C. *Vista Astr.* **14**, 163-289 (1972).
- Cardwell, J. A. R. & Coulson, I. M. *Mon. Not. R. astr. Soc.* **212**, 879-888 (1985).
- Cardwell, J. A. R. & Coulson, I. M. preprint. South African Astronomical Observatory.
- Mould, J. R., Da Costa, G. S. & Crawford, M. D. *Astrophys. J.* **280**, 595-599 (1984).
- Schommer, R. A., Olszewski, E. W. & Aaronson, M. *Astrophys. J.* **285**, L53-L57 (1984).
- Hodge, P. *Astrophys. J.* **264**, 470-475 (1983).
- Westerlund, B. E. *Mon. Not. R. astr. Soc.* **127**, 429-448 (1964).

C_{60} : Buckminsterfullerene

H. W. Kroto*, J. R. Heath, S. C. O'Brien, R. F. Curl
& R. E. Smalley

Rice Quantum Institute and Departments of Chemistry and Electrical Engineering, Rice University, Houston, Texas 77251, USA

During experiments aimed at understanding the mechanisms by which long-chain carbon molecules are formed in interstellar space and circumstellar shells¹, graphite has been vaporized by laser irradiation, producing a remarkably stable cluster consisting of 60 carbon atoms. Concerning the question of what kind of 60-carbon atom structure might give rise to a superstable species, we suggest a truncated icosahedron, a polygon with 60 vertices and 32 faces, 12 of which are pentagonal and 20 hexagonal. This object is commonly encountered as the football shown in Fig. 1. The C_{60} molecule which results when a carbon atom is placed at each vertex of this structure has all valences satisfied by two single bonds and one double bond, has many resonance structures, and appears to be aromatic.

The technique used to produce and detect this unusual molecule involves the vaporization of carbon species from the surface of a solid disk of graphite into a high-density helium flow, using a focused pulsed laser. The vaporization laser was the second harmonic of Q-switched Nd:YAG producing pulse energies of ~30 mJ. The resulting carbon clusters were expanded in a supersonic molecular beam, photoionized using an excimer laser, and detected by time-of-flight mass spectrometry. The vaporization chamber is shown in Fig. 2. In the experiment the pulsed valve was opened first and then the vaporization laser was fired after a precisely controlled delay. Carbon species were vaporized into the helium stream, cooled and partially equilibrated in the expansion, and travelled in the resulting molecular beam to the ionization region. The clusters were ionized by direct one-photon excitation with a carefully synchronized excimer laser pulse. The apparatus has been fully described previously²⁻⁵.

The vaporization of carbon has been studied previously in a very similar apparatus⁶. In that work clusters of up to 190 carbon atoms were observed and it was noted that for clusters of more than 40 atoms, only those containing an even number of atoms were observed. In the mass spectra displayed in ref. 6, the C_{60} peak is the largest for cluster sizes of >40 atoms, but it is not completely dominant. We have recently re-examined this system and found that under certain clustering conditions the C_{60} peak can be made about 40 times larger than neighbouring clusters.

Figure 3 shows a series of cluster distributions resulting from variations in the vaporization conditions evolving from a cluster distribution similar to that observed in ref. 3, to one in which C_{60} is totally dominant. In Fig. 3c, where the firing of the vaporization laser was delayed until most of the He pulse had passed, a roughly gaussian distribution of large, even-numbered clusters with 38–120 atoms resulted. The C_{60} peak was largest but not dominant. In Fig. 3b, the vaporization laser was fired at the time of maximum helium density; the C_{60} peak grew into a feature perhaps five times stronger than its neighbours, with the exception of C_{70} . In Fig. 3a, the conditions were similar to those in Fig. 3b but in addition the integrating cup depicted in Fig. 2 was added to increase the time between vaporization and expansion. The resulting cluster distribution is completely dominated by C_{60} , in fact more than 50% of the total large cluster abundance is accounted for by C_{60} ; the C_{70} peak has diminished in relative intensity compared with C_{60} , but remains rather prominent, accounting for ~5% of the large cluster population.

Our rationalization of these results is that in the laser vaporization, fragments are torn from the surface as pieces of the planar

Fig. 1 A football (in the United States, a soccerball) on Texas grass. The C_{60} molecule featured in this letter is suggested to have the truncated icosahedral structure formed by replacing each vertex on the seams of such a ball by a carbon atom.



graphite fused six-membered ring structure. We believe that the distribution in Fig. 3c is fairly representative of the nascent distribution of larger ring fragments. When these hot ring clusters are left in contact with high-density helium, the clusters equilibrate by two- and three-body collisions towards the most stable species, which appears to be a unique cluster containing 60 atoms.

When one thinks in terms of the many fused-ring isomers with unsatisfied valences at the edges that would naturally arise from a graphite fragmentation, this result seems impossible: there is not much to choose between such isomers in terms of stability. If one tries to shift to a tetrahedral diamond structure, the entire surface of the cluster will be covered with unsatisfied valences. Thus a search was made for some other plausible structure which would satisfy all sp^2 valences. Only a spheroidal structure appears likely to satisfy this criterion, and thus Buckminster Fuller's studies were consulted (see, for example, ref. 7). An unusually beautiful (and probably unique) choice is the truncated icosahedron depicted in Fig. 1. As mentioned above, all valences are satisfied with this structure, and the molecule appears to be aromatic. The structure has the symmetry of the icosahedral group. The inner and outer surfaces are covered with a sea of π electrons. The diameter of this C_{60} molecule is ~7 Å, providing an inner cavity which appears to be capable of holding a variety of atoms⁸.

Assuming that our somewhat speculative structure is correct, there are a number of important ramifications arising from the existence of such a species. Because of its stability when formed under the most violent conditions, it may be widely distributed in the Universe. For example, it may be a major constituent of circumstellar shells with high carbon content. It is a feasible constituent of interstellar dust and a possible major site for

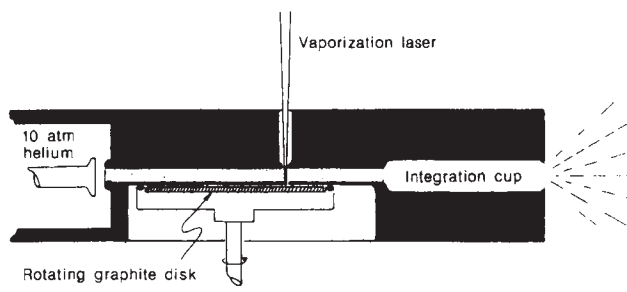


Fig. 2 Schematic diagram of the pulsed supersonic nozzle used to generate carbon cluster beams. The integrating cup can be removed at the indicated line. The vaporization laser beam (30–40 mJ at 532 nm in a 5-ns pulse) is focused through the nozzle, striking a graphite disk which is rotated slowly to produce a smooth vaporization surface. The pulsed nozzle passes high-density helium over this vaporization zone. This helium carrier gas provides the thermalizing collisions necessary to cool, react and cluster the species in the vaporized graphite plasma, and the wind necessary to carry the cluster products through the remainder of the nozzle. Free expansion of this cluster-laden gas at the end of the nozzle forms a supersonic beam which is probed 1.3 m downstream with a time-of-flight mass spectrometer.

* Permanent address: School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

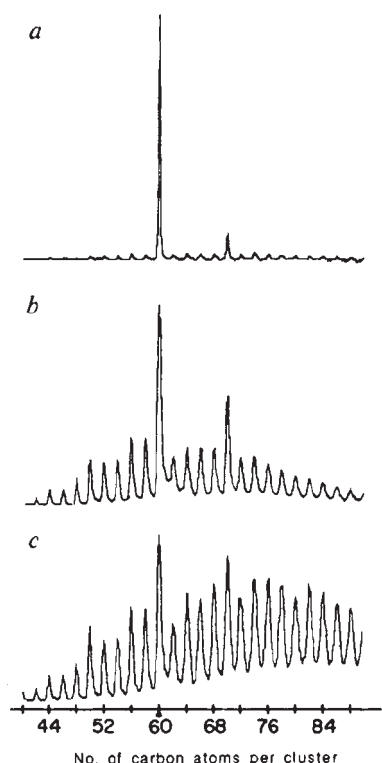


Fig. 3 Time-of-flight mass spectra of carbon clusters prepared by laser vaporization of graphite and cooled in a supersonic beam. Ionization was effected by direct one-photon excitation with an ArF excimer laser (6.4 eV , 1 mJ cm^{-2}). The three spectra shown differ in the extent of helium collisions occurring in the supersonic nozzle. In *c*, the effective helium density over the graphite target was less than 10 torr —the observed cluster distribution here is believed to be due simply to pieces of the graphite sheet ejected in the primary vaporization process. The spectrum in *b* was obtained when roughly 760 torr helium was present over the graphite target at the time of laser vaporization. The enhancement of C_{60} and C_{70} is believed to be due to gas-phase reactions at these higher clustering conditions. The spectrum in *a* was obtained by maximizing these cluster thermalization and cluster-cluster reactions in the 'integration cup' shown in Fig. 2. The concentration of cluster species in the especially stable C_{60} form is the prime experimental observation of this study.

surface-catalysed chemical processes which lead to the formation of interstellar molecules. Even more speculatively, C_{60} or a derivative might be the carrier of the diffuse interstellar lines⁹.

If a large-scale synthetic route to this C_{60} species can be found, the chemical and practical value of the substance may prove extremely high. One can readily conceive of C_{60} derivatives of many kinds—such as C_{60} transition metal compounds, for example, $C_{60}\text{Fe}$ or halogenated species like $C_{60}\text{F}_{60}$ which might be a super-lubricant. We also have evidence that an atom (such as lanthanum⁸ and oxygen¹) can be placed in the interior, producing molecules which may exhibit unusual properties. For example, the chemical shift in the NMR of the central atom should be remarkable because of the ring currents. If stable in macroscopic, condensed phases, this C_{60} species would provide a topologically novel aromatic nucleus for new branches of organic and inorganic chemistry. Finally, this especially stable and symmetrical carbon structure provides a possible catalyst and/or intermediate to be considered in modelling prebiotic chemistry.

We are disturbed at the number of letters and syllables in the rather fanciful but highly appropriate name we have chosen in the title to refer to this C_{60} species. For such a unique and centrally important molecular structure, a more concise name would be useful. A number of alternatives come to mind (for

example, ballene, sphereene, soccerene, carbosoccer), but we prefer to let this issue of nomenclature be settled by consensus.

We thank Frank Tittel, Y. Liu and Q. Zhang for helpful discussions, encouragement and technical support. This research was supported by the Army Research Office and the Robert A. Welch Foundation, and used a laser and molecular beam apparatus supported by the NSF and the US Department of Energy. H.W.K. acknowledges travel support provided by SERC, UK. J.R.H. and S.C.O.B. are Robert A. Welch Predoc-toral Fellows.

Received 13 September; accepted 18 October 1985.

1. Heath, J. R. *et al. Astrophys. J.* (submitted).
2. Dietz, T. G., Duncan, M. A., Powers, D. E. & Smalley, R. E. *J. chem. Phys.* **74**, 6511–6512 (1981).
3. Powers, D. E. *et al. J. phys. Chem.* **86**, 2556–2560 (1982).
4. Hopkins, J. B., Langridge-Smith, P. R. R., Morse, M. D. & Smalley, R. E. *J. chem. Phys.* **78**, 1627–1637 (1983).
5. O'Brien, S. C. *et al. J. chem. Phys.* (submitted).
6. Rohlfing, E. A., Cox, D. M. & Kaldor, A. *J. chem. Phys.* **81**, 3322–3330 (1984).
7. Marks, R. W. *The Dymaxion World of Buckminster Fuller* (Reinhold, New York, 1960).
8. Heath, J. R. *et al. J. Am. chem. Soc.* (in the press).
9. Herbig, E. *Astrophys. J.* **196**, 129–160 (1975).

High-resolution solid-state NMR of quadrupolar nuclei

Eric Oldfield, Hye Kyung C. Timken,
Ben Montez & R. Ramachandran

School of Chemical Sciences, University of Illinois at
Urbana-Champaign, 505 South Mathews Avenue, Urbana,
Illinois 61801, USA

Quadrupolar nuclei are the most abundant nuclear magnetic resonance (NMR)-receptive nuclei in the Earth's crust, and in many amorphous materials of technological interest (such as zeolite catalysts, ceramics and alloys), and have thus been intensively studied^{1–7}. Of particular interest is the ability to resolve and quantitate the various types of sites present in a given material. Here we present a very simple, yet we believe powerful, approach towards the resolution of chemically non-equivalent sites in solids, which combines a conventional high-field spin-echo NMR method with the resolution enhancement of the 'quadrupole shift' approach^{4,7}. We demonstrate its application to the complete resolution of both the $(1/2, 3/2)$ and $(3/2, 5/2)$ transitions of the ^{27}Al nuclei in a mixture of potassium and ammonium alums ($\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$ and $\text{NH}_4\text{Al}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$).

The major problem in spin-echo studies of most quadrupolar nuclei in solids is that a series of echoes is usually obtained^{8,9}. Solomon first reported⁸ the refocusing of first-order quadrupolar interactions using two radio-frequency pulses of the same phase, and Butterworth⁹ was able to differentiate quadrupolar from purely magnetic interactions in various alloys. However, this approach is not particularly useful for minerals and other non-metallic materials, where large magnetic inhomogeneities are absent.

Fortunately, however, Bonera and Galimberti¹⁰ and Weisman and Bennett¹¹ showed that considerable enhancement of the 'desired' 2τ echo could be achieved by introducing a 90° phase-shift between the two radio-frequency pulses. Moreover, Weisman and Bennett showed (for nuclear spin $I = 5/2$) for the special case of a $90^\circ - \tau - 45^\circ$ pulse sequence that the amplitude of the 2τ echo (central plus quadrupolar contribution) was maximized, while the 'allowed' $3/2\tau$ plus 3τ echo amplitude was minimized¹¹. In practice, the 2τ echo is often the only one observed¹¹, and we show here that the Fourier transform of the echo can yield an essentially undistorted $(1/2, -1/2; 1/2, 3/2; 3/2, 5/2)$ powder spectrum.

Figure 1*a* and *b* show the ^{23}Na and ^{27}Al spin-echo spectra of NaNO_3 and $\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$, obtained at magnetic field strengths of 8.45 and 3.52 T, respectively, the response of the potassium alum being obtained in the presence of ^1H dipolar

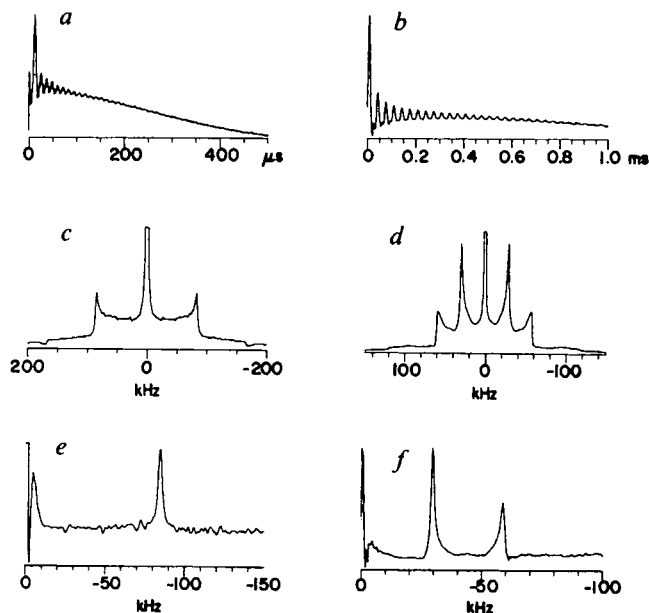


Fig. 1 Quadrupole-echo NMR spectra of ^{23}Na in NaNO_3 and ^{27}Al in $\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$, obtained at magnetic field strengths of 8.45 and 3.52 T, respectively. *a*, ^{23}Na spin-echo from NaNO_3 ; *b*, ^{27}Al spin-echo from $\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$; *c*, Fourier transform of half-echo in *a* showing central (1/2, -1/2) and satellite (1/2, 3/2) transitions; *d*, Fourier transform of half-echo in *b* showing central (1/2, -1/2) and satellite (1/2, 3/2; 3/2, 5/2) transitions; *e*, de-Paked version of *c* showing the 90° edge, corresponding to $e^2qQ/h \sim 339$ kHz, η assumed ~ 0 ; *f*, de-Paked version of *d* showing the 90° edges of the (1/2, 3/2) and (3/2, 5/2) transitions, corresponding to $e^2qQ/h \sim 395$ kHz, η assumed ~ 0 .

decoupling (at 150 MHz). No subsidiary echoes were observed for ^{27}Al and none were expected for ^{23}Na , even though, with our present instrumentation, we were unable to use pulses much stronger than the overall spectral breadth. Nevertheless, Fourier transformation of both responses yielded conventional frequency domain powder-pattern spectra (Fig. 1*c, d*) in which the 90° edges of the 1/2, 3/2 (Na, Al) and 3/2, 5/2 (Al) transitions were well resolved.

To provide a more convenient 'high-resolution' presentation, we have applied the 'de-Pakeing' algorithm developed by Bloom *et al.*^{12,13} to the non-integral spin spectra of Fig. 1*c, d*, and show the results in Fig. 1*e* and *f*. For both systems, there were no appreciable second-order effects, and asymmetry parameters $\eta \sim 0$, so narrow-line spectra corresponding to nuclear quadrupole coupling constants, $e^2qQ/h = 339$ kHz (NaNO_3) and 395 kHz ($\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$) were obtained, in good agreement with previous single crystal determinations¹⁴⁻¹⁷.

To illustrate further the usefulness of the method, which can be implemented on a conventional high-field NMR instrument, Fig. 2*a-c* gives the spin-echo, spin-echo Fourier transform powder pattern and de-Paked lineshape for an equimolar mixture of potassium and ammonium alums, again obtained in the presence of ^1H dipolar decoupling. The K^+ and NH_4^+ alum signals have their expected 1:1 intensity ratios. In both cases, the potassium and ammonium alum signals are fully resolved, the outer satellite transitions being slightly more resolved than the inner ones (even though their component linewidths are greater). Further line-narrowing may be possible in other cases by application of homonuclear line-narrowing techniques (when e^2qQ/h values are very small). The quadrupole coupling constants (395 kHz and 441 kHz for potassium alum and ammonium alum, respectively) and asymmetry parameters ($\eta = 0$ for both) show excellent agreement with the single crystal NMR data¹⁵⁻¹⁷, which give $e^2qQ/h = 400$ kHz for potassium alum and 446 kHz for ammonium alum, and $\eta = 0$ for both (since the point symmetry of the ^{27}Al is $\bar{3}$ in both structures).

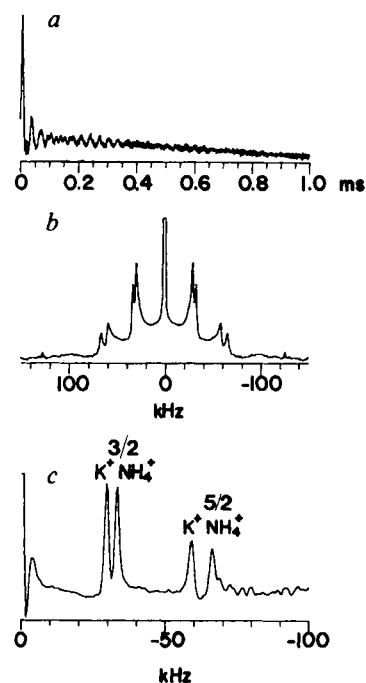


Fig. 2 Quadrupole-echo NMR spectra of ^{27}Al in a 1:1 molar mixture of $\text{KAl}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$ and $\text{NH}_4\text{Al}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$, obtained at a magnetic field strength of 3.52 T, in the presence of ^1H dipolar decoupling. *a*, ^{27}Al spin-echo; *b*, Fourier transform of half-echo showing central (1/2, -1/2) and overlapping satellite (1/2, 3/2; 3/2, 5/2) transitions for the K^+ and NH_4^+ alums. *c*, de-Paked version of *b* showing complete resolution of both K^+ and NH_4^+ alums in both satellite transitions, corresponding to $e^2qQ/h \sim 395$, 441 kHz, η assumed ~ 0 .

Overall, we believe that the above results indicate considerable promise for high-field NMR techniques in the differentiation of signals in powder samples from quadrupolar nuclei of identical chemical shifts. The technique should be applicable to systems having larger e^2qQ/h values by using high-power wide-band transmitters at higher field (using smaller sample, low- Q probes). Such an application was limited in the present case by the use of pulses having ~ 0.5 - μs rise and fall times, but should not be limited to systems having long spin-lattice relaxation times. It can be readily implemented (at least for small e^2qQ/h values) on existing conventional Fourier transform NMR instruments, and could thus be extremely useful in the structural analysis of a variety of amorphous solids of interest in the earth, materials and biological sciences. For all nuclei, operation with very wide-bore magnets (to house the very high-voltage capacitors needed for short pulse widths) at the highest magnetic field strengths possible (to minimize second-order quadrupole interactions and 'ringing' effects) will, however, be essential in order to maximize the usefulness of this approach.

We thank M. Bloom and H. C. Jarrell for providing their spectral de-Pakeing programs. This work was supported by the US NSF Solid-State Chemistry Program and by the Biophysics Program, and in part by the US Department of Energy (to H.K.C.T.).

Received 6 June; accepted 10 September 1985.

1. Fyfe, C. A., Thomas, J. M., Klinowski, J. & Gobbi, G. C. *Angew. Chem. Int. Edn. Engl.* **22**, 259-275 (1983).
2. Oldfield, E. & Kirkpatrick, R. J. *Science* **227**, 1537-1544 (1985).
3. Ganapathy, S., Schramm, S. & Oldfield, E. J. *chem. Phys.* **77**, 4360-4365 (1982).
4. Schramm, S. & Oldfield, E. *JCS chem. Commun.* 980-981 (1982).
5. Weitekamp, D. P., Bielecki, A., Zax, D., Zilm, K. & Pines, A. *Phys. Rev. Lett.* **50**, 1807-1810 (1983).
6. Bielecki, A. *et al. J. chem. Phys.* **80**, 2232-2234 (1984).
7. Zax, D. B., Bielecki, A., Pines, A. & Sinton, S. W. *Nature* **312**, 351-352 (1984).
8. Solomon, I. *Phys. Rev.* **110**, 61-65 (1958).
9. Butterworth, J. *Proc. Phys. Soc.* **86**, 297-304 (1965).
10. Bonera, G. & Galimberti, M. *Solid St. Commun.* **4**, 589-591 (1966).
11. Weisman, I. D. & Bennett, L. H. *Phys. Rev.* **181**, 1341-1350 (1969).

12. Bloom, M., Davis, J. H. & MacKay, A. L. *Chem. Phys. Lett.* **80**, 198–202 (1981).
 13. Sternin, E., Bloom, M. & MacKay, A. L. *J. magn. Res.* **55**, 274–282 (1983).
 14. Pound, R. V. *Phys. Rev.* **79**, 685–702 (1950).
 15. Burns, G. J. *chem. Phys.* **32**, 1585–1586 (1960).
 16. Weiden, N. & Weiss, A. *Ber. Bunsenges. phys. Chem.* **78**, 1031–1050 (1974).
 17. Weiden, N. & Weiss, A. *Ber. Bunsenges. phys. Chem.* **79**, 557–563 (1975).

A hydrated aluminophosphate with both 4.8² and 6³ sheets in the 4-connected framework

Joseph J. Pluth & Joseph V. Smith

Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

Porous aluminosilicates of the zeolite type have become important industrially as ion-exchangers, molecular sieves and catalysts^{1,2}. Topologically, the tetrahedral centres of the oxygen tetrahedra lie at the nodes of various 4-connected three-dimensional nets. Four-connected porous frameworks (AlPO₄) composed of alternating AlO₄ and PO₄ tetrahedra are the basis of a new family of synthetic molecular sieves, which may become important industrially because they lack the strong ion-exchange and catalytic properties of the zeolites³. Aluminophosphates display a wide range of crystal structures because Al can be coordinated to either four, five or six oxygen atoms, some of which may belong to either hydroxyl or water species^{3–12}. As part of a systematic study of the crystal chemistry of aluminophosphates, we have found that the framework of synthetic-phase H3 (ref. 13) (AlPO₄·1.5 H₂O) contains PO₄ tetrahedra alternating between AlO₄ tetrahedra and AlO₄(H₂O)₂ octahedra. As reported here, the 4-connected three-dimensional net formed by linking adjacent Al and P atoms is a new type containing 6.6.6 and 4.8.8 two-dimensional nets joined by up-down linkages. This coexistence of two types of two-dimensional nets in a three-dimensional net suggests that the principle of parsimony should not be adopted too strictly in attempts to invent new nets of potential relevance to molecular sieve technology.

Crystals of a hydrated aluminium phosphate were synthesized by S. T. Wilson, B. M. Lok and E. M. Flanigen of Union Carbide Corporation. The X-ray powder diffraction pattern of the crystals matches that in Fig. 2 of ref. 13. Full details of the X-ray structure determination will be published elsewhere¹². Cell dimensions and atomic coordinates (Table 1) were refined by the least-squares method. Refinement was straightforward except for the H atoms. The positions of H-1 and H-2 were constrained to lie at 0.85 Å from O-9; the same was true for O-10, H-3 and H-4. Whereas these two water oxygens are bonded to Al-2, the last oxygen O-11, is not part of the framework and its position is weakly constrained. Because hydrogen atoms could not be located even from a constrained refinement, it is assumed that O-11 is part of a loosely bonded water molecule.

Figure 1 is a stereo plot showing the positions of all located atoms. Each P atom is tetrahedrally coordinated, and shares an oxygen atom with four adjacent Al atoms. Similarly, each Al-1 atom shares an oxygen atom with one P-2 and three P-1 atoms. Each Al-2 atom, however, is bonded to two water molecules as well as four framework oxygens. Such an octahedral complex, AlO₄(H₂O)₂, is also found in metavariscite⁵. The atomic positions in Table 1 correspond to AlPO₄·1.5 H₂O instead of the AlPO₄·1.67 H₂O listed by d'Yvoire¹³. The complex dehydration and rehydration properties of H₃ could be interpreted in terms of easy loss of the water molecule at position O-11, conversion of the octahedral complex AlO₄(OH)₂ to a tetrahedral complex AlO₄ by loss of two water molecules at a higher temperature, and adsorption of water by the dehydration product on cooling in a humid atmosphere; however, further structure determinations are needed to clarify this.

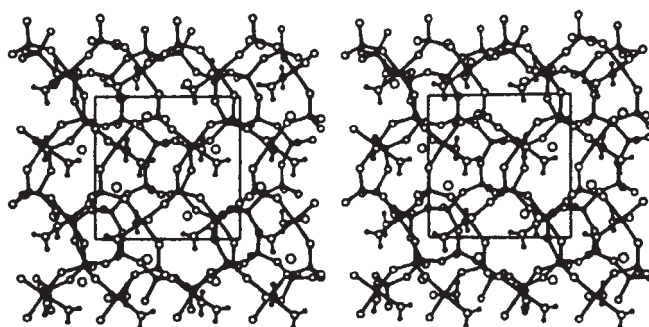


Fig. 1 Stereo plot of H3 structure type, projected down the *a* axis with *c* pointing upwards and *b* sideways. Small open and solid circles at line intersections represent Al and P atoms, respectively. Large open circles represent oxygen atoms. Pairs of H atoms (small dots) are shown for each water molecule attached to an Al atom, but not for the loosely bonded water molecule.

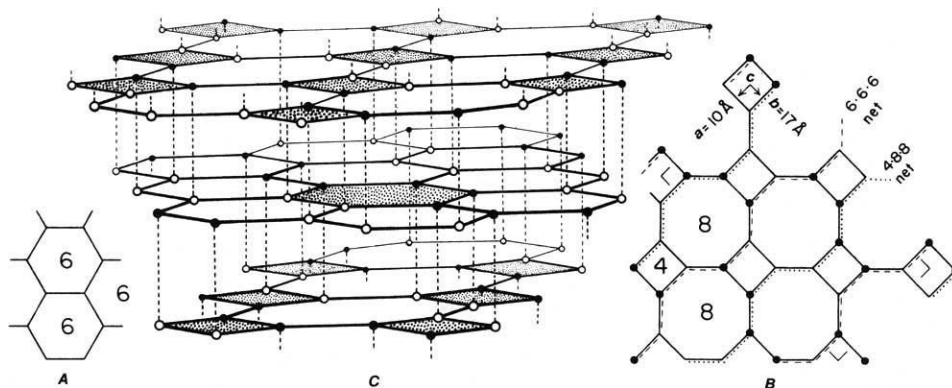
When the water molecules are ignored, the structure can be regarded as a 4-connected tetrahedral framework which can be compared with the growing list of theoretical and three-dimensional nets (ref. 14 and refs therein; also refs 15, 16). The H3 net is novel, and it is interesting because it is based on up-down linkages between both 6³ and 4.8² nets. A portion of a 3-connected 6.6.6 two-dimensional net is shown in plan view in Fig. 2A; each node is connected by three edges, each of which is part of a circuit of six edges. The continuous lines of Fig. 2B represent part of a 3-connected 4.8.8 two-dimensional net. An infinite number of three-dimensional nets can be obtained by stacking horizontal 6.6.6 two-dimensional nets in a vertical pile and connecting adjacent nodes by a vertical linkage¹⁷. Because only one linkage can be added to each 3-connected node, there are two choices for each node. The simplest three-dimensional nets have been enumerated, and some are found in crystalline materials¹⁷. Similar enumeration has been done for 4.8.8 nets connected by up-down linkages¹⁸. The H3 tetrahedral framework is shown in clinographic perspective in Fig. 2C. A 6.6.6 net (centre) is connected to 4.8.8 nets above and below it by vertical linkages (dashed lines). Each 3-connected node marked by an open circle is joined to a node above it (solid circle) so that each node becomes 4-connected. Regular alternation of 6.6.6 and 4.8.8 nets in the vertical stack generates an infinite 4-connected three-dimensional 3D net. The 6.6.6 and

Table 1 The atomic positions ($\times 10^5$) and mean-square displacements ($\text{\AA}^2$) of AlPO₄-H3

Atom	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>	<i>U</i> _{eq} or <i>U</i> _{iso}
P(1)	44,769 (4)	34,071 (8)	13,498 (9)	84 (2)
P(2)	29,050 (4)	6,565 (8)	33,879 (9)	78 (2)
Al(1)	44,808 (4)	8,470 (9)	32,419 (10)	89 (3)
Al(2)	28,207 (5)	33,224 (9)	14,654 (10)	79 (3)
O(1)	37,433 (10)	37,652 (22)	17,006 (24)	146 (7)
O(2)	46,047 (12)	16,616 (22)	48,037 (24)	174 (7)
O(3)	49,520 (10)	44,967 (21)	19,709 (24)	163 (7)
O(4)	46,889 (11)	20,038 (20)	19,416 (23)	134 (7)
O(5)	27,397 (10)	5,819 (20)	49,128 (22)	114 (7)
O(6)	36,522 (10)	1,895 (22)	31,355 (24)	160 (7)
O(7)	28,329 (11)	21,582 (20)	29,532 (23)	133 (7)
O(8)	25,855 (10)	47,583 (20)	25,919 (22)	112 (7)
O(9)	18,446 (12)	28,311 (27)	11,931 (31)	220 (8)
O(10)	30,512 (12)	17,298 (23)	3,398 (26)	155 (8)
O(11)	11,929 (14)	14,398 (33)	36,268 (37)	568 (13)
H(1)	15,752 (12)	29,879 (27)	5,232 (31)	197 (37)
H(2)	16,096 (12)	25,744 (27)	18,827 (31)	150 (28)
H(3)	28,123 (12)	10,248 (23)	5,406 (26)	66 (16)
H(4)	30,108 (12)	19,016 (23)	-5,101 (26)	54 (17)

a = 19.352(1), *b* = 0.727(1), *c* = 9.762(1) Å; orthorhombic, Pbc_a. Room temperature. CuK_α. Crystal 40 × 40 × 220 μm, elongated c. 1,439 unique diffraction intensities. Anisotropic thermal parameters. *R* = 0.04. *U*_{iso}($\times 10^3$) is given for hydrogen atoms, and *U*_{eq}[(=10³/3)Σ*i*Σ*j**U*_{*ij*}*a*_{*i*}²*a*_{*j*}²(*a*_{*i*}*a*_{*j*})] is given for the other atoms. Number in parentheses represent standard error.

Fig. 2 Topological drawings. **A**, Plan view of a portion of a 6.6.6 3-connected two-dimensional net. **B**, Plan view of a portion of a 4.8.8 3-connected two-dimensional net (continuous lines). The corner marks outline the *c*-axis projection of the unit cell for the H3 structure type idealized to the regular geometry for up and down linkages from a 4.8.8 two-dimensional net (up, node marked by solid circle; down, unmarked node). Dotted and broken lines show, respectively, side views of the 4.8.8 and 6.6.6 nets in *C*. **C**, Perspective drawing of the up-down linkages (dashed lines) between alternating 4.8.8 and 6.6.6 two-dimensional nets. Twelve 4-rings and one 6-ring are stippled to strengthen the perspective.



4.8.8 nets are distorted from the ideal shapes when the linkages are vertical. In the H3 structure (Fig. 1), the nets are considerably crinkled to allow for bonding to the water molecules, but may become more regular on complete dehydration.

The H3 structure can also be described in terms of up-down linkages between a vertical stack of 4.8.8 nets. Figure 2B shows a plan view of a three-dimensional net obtained when adjacent groups of four linkages (solid circles) alternate with four down-down linkages (unmarked nodes). The unit cell for the most regular geometry and an edge length of 3.1 Å is orthorhombic, with $a = 17$ Å, $b = 10$ Å, $c = 8.5$ Å, space group *Abmm*, 32 nodes (Z_c) per cell, 16 nodes (Z_i) in the asymmetric unit, circuit symbol $4^3 6^2 8$. (Z_c is the number of atoms in the crystallographic unit cell and Z_i is the number of atoms in the smallest unit cell.) This net was not enumerated previously¹⁸ because its a repeat of 17 Å is longer than the upper limit of 15 Å selected by Smith and Rinaldi¹⁹; it is now denoted as net 24a. Crinkling of the framework in the H3 structure changes the above cell to $a = 19.35$ Å, $b = 9.73$ Å, $c = 9.76$ Å, *Pbca*; in particular, the double-crankshaft chain along the *c* axis is stretched out considerably over that in the feldspar structure. It should be emphasized that there are two 4.8.8 two-dimensional nets in the H3 framework; one is seen in plan view in Fig. 2B, and the other in side view (dotted lines). The 6.6.6 net is also shown in side view (dashed lines).

The complex bonding properties of the Al and P atoms in the H3 structure, and in particular the existence of alternating 6.6.6 and 4.8.8 two-dimensional nets, suggests that the principle of parsimony should not be taken too strictly in attempts to invent new nets and crystalline materials of potential use in molecular sieve technology. Furthermore, the geometrical flexibility of the two-dimensional component nets of three-dimensional nets (for example, 6^3 net; ref. 20) must be considered in attempts to invent new types of linkages in three-dimensional nets.

We thank E. M. Flanigen, L. Patton and S. T. Wilson for reviewing the manuscript, the NSF for grant CHE 84-05167, Union Carbide Corporation (Linde Molecular Sieves) for financial support, and three referees for comments which led to replacement of an incomprehensible projection by a perspective drawing.

Received 9 April; accepted 23 September 1985.

1. Breck, D. W. *Molecular Sieves* (Wiley, New York, 1974).
2. Rabo, J. A. (ed.) *Am. chem. Soc. Monogr.* 171 (1976).
3. Wilson, S. T., Lok, B. M., Messina, C. A., Cannan, T. R. & Flanigen, E. M. *J. Am. chem. Soc.* **104**, 1146-1147 (1982).
4. Schwarzenbach, D. Z. *Kristallogr.* **123**, 161-185 (1966).
5. Kneip, R. & Mootz, D. *Acta crystallogr.* **B29**, 2292-2294 (1973).
6. Kneip, R., Mootz, D. & Vegas, A. *Acta crystallogr.* **B33**, 263-265 (1977).
7. Keegan, T. D., Araki, T. & Moore, P. B. *Am. Miner.* **64**, 1243-1247 (1979).
8. Bennett, J. M., Cohen, J. P., Flanigen, E. M., Pluth, J. J. & Smith, J. V. *Am. chem. Soc. Symp. Ser.* **218**, 109-118 (1983).
9. Pluth, J. J., Smith, J. V., Bennett, J. M. & Cohen, J. P. *Acta crystallogr.* **C40**, 2008-2011 (1984).
10. Bennett, J. M., Cohen, J. P., Artioli, G., Pluth, J. J. & Smith, J. V. *Inorg. Chem.* **24**, 188-193 (1985).

11. Parise, J. B. *JCS Chem. Commun.* 1449-1450 (1984).
12. Pluth, J. J., Smith, J. V. & Bennett, J. M. *Acta crystallogr.* *C* (submitted).
13. d'Yvoire, F. *Bull. Soc. chim. Fr.* 1762-1776 (1962).
14. Smith, J. V. Z. *Kristallogr.* **163**, 191-198 (1983).
15. Meier, W. M. & Olson, D. H. *Atlas of Zeolite Structure Types* (Polycrystal Book Service, Pittsburgh, Pennsylvania, 1978).
16. Meier, W. M. & Moeck, H. J. *J. Solid State Chem.* **24**, 349-355 (1979).
17. Smith, J. V. *Am. Miner.* **62**, 703-709 (1977).
18. Smith, J. V. *Am. Miner.* **63**, 960-969 (1978).
19. Smith, J. V. & Rinaldi, F. *Mineralog. Mag.* **33**, 202-212 (1962).
20. Meier, W. M. *Natural Zeolites* (eds Sand, L. B. & Mumpton, F. A.) 99-103 (Pergamon, Oxford, 1978).

Meteoritic evidence that graphite is rare in the interstellar medium

Joseph A. Nuth

Solar System Exploration Division, Office of Space Science and Applications, NASA Headquarters, Washington, DC 20546, USA

This letter discusses the carbonaceous material from several distinct astrophysical environments that has been identified in primitive meteorites. Most of this material is not well-crystallized graphite, but ranges from kerogen-like macromolecular organic matter to poorly graphitized carbon. Of the small fraction of the carbonaceous material which is graphite, most of this may have resulted from the graphitization of macromolecular precursors in the solar nebula. Since graphite is more stable than those precursors, pre-existing interstellar graphite should have survived in meteorites to at least the same extent as did other forms of carbon. I conclude from this dearth of graphite that graphitic carbon was not a major component of the interstellar dust at the time of the formation of the Solar System.

Several forms of carbon are present in the primitive meteorites¹. Some of this carbon may have been produced in the primitive solar nebula¹, some by ion-molecule reactions in interstellar space^{2,3} and a small fraction could have been formed in the expanding shells of red giants⁴, novae⁵ and supernovae⁶. The bulk of meteoritic carbon is 'ordinary' in the sense that it is neither the carrier of an isotopically distinct rare gas component nor itself distinguishable by an 'anomalous' $^{13}\text{C}/^{12}\text{C}$ ratio or associated anomalous $^{15}\text{N}/^{14}\text{N}$ or D/H ratios. The 'exotic' component of primitive meteorites is important not only because of its potential to record subtle clues about the nucleosynthetic process⁷, but also as evidence that some interstellar materials survived the formation of the Solar System and were incorporated relatively intact into meteorites, a possibility suggested by Cameron⁸ as early as 1973.

Anomalous material is that in which the isotopic ratio of an element differs from the standard 'cosmic' isotopic ratio. Both refractory and volatile elements have been shown to be isotopi-

cally anomalous (reviewed in ref. 7). Anomalies in volatile elements are most relevant in the present context because their preservation indicates that some delicate features of pre-solar interstellar grains can still be found in primitive meteorites. Numerous isotopically anomalous rare gas components are known to exist in the meteorites and several of these are thought to pre-date the collapse of the solar nebula. A ^{22}Ne -rich gas component is released from C1 and C2 chondrites at temperatures of 1,075–1,375 K (ref. 9). Recent work¹⁰ has shown that pure ^{22}Ne can be released from a carbonaceous carrier phase at temperatures as low as 875 K. It is now thought that this neon results from the decay of ^{22}Na which condensed on grains formed in the ejecta of a nova. A mixture of xenon isotopes, designated xenon HL, is released from carbonaceous chondrites at temperatures of 875–1,375 K (ref. 11). Xenon HL is enriched in the xenon isotopes 124, 126, 134 and 136 and is thought to have formed in the outflow around a supernova. Another mixture of xenon isotopes, released at 1,375–1,875 K, is enriched in the xenon isotopes 128, 130 and 132 (ref. 4). The composition of this component matches the predicted ratio of xenon isotopes produced via the s-process in red giant stars¹².

It is impossible to overstate the potential influence of secondary metamorphic processes which may have altered the structure or composition of interstellar grains that had survived the 'nebular' phase of Solar System formation. These processes are undoubtedly the source of the abundant graphite found in iron meteorites¹³ and it has been suggested that similar processes may be the source of all meteoritic graphite, even that found in chondrites¹⁴. The existence of the anomalies discussed above, however, indicates that it is possible for some pre-solar grains to have survived relatively intact to the present time. The observed release temperatures indicate that many surviving pre-solar materials were never exposed to temperatures above 1,000 K. Such conditions should not be unfavourable enough to destroy pre-existing interstellar graphite grains and, as noted previously, it has been suggested that the graphite observed in chondrites may have formed in this environment¹⁴.

Most accepted models of interstellar grains consist of various mixtures of graphite, 'silicate' and additional minor components^{15,16}. Graphite is included in these models not only because it can explain the observed peak in the interstellar extinction curve at 220 nm, given certain specific conditions, but also because it is the most stable form of solid carbon in interstellar conditions. (It has the additional advantage that a well-determined set of optical constants is available from the literature.) Millar and Duley¹⁷ have suggested that the 220-nm 'bump' is produced by crystalline magnesium oxide grains, while Hoyle and co-workers¹⁸ have suggested that the observed extinction feature might be caused by various organic materials found in biological systems. More recently, Sakata *et al.*¹⁹ have shown that a 'quenched carbonaceous composite' (QCC), which is a solid soot-like residue produced in a methane plasma discharge, might yield an acceptable fit to the 220-nm extinction peak. Finally, ultraviolet observations by Hecht *et al.*²⁰ have shown that the extinction properties of the dust produced around R Corona Borealis stars are consistent with the properties of amorphous or glassy carbon²¹, rather than with those of graphite. This is especially significant as R Corona Borealis stars are both carbon-rich and hydrogen-deficient ($\text{C}/\text{H} > 100$ in some cases). In such conditions, hydrocarbon formation should be greatly suppressed and there would be relatively few kinetic inhibitions to the formation of graphite²². Failure to form graphite particles in such favourable conditions seems to lessen the probability of their formation in the more common, hydrogen-rich stars that are the source of most interstellar grains²³.

True 'interstellar' carbon has been detected in meteorites² and is distinguished by a greatly enhanced D/H ratio, thought to have been produced through ion-molecule reactions in very cold interstellar clouds²⁴. Such material is not graphite, but consists of a macromolecular kerogen-like polymer¹, a typical composition for which is 76.5% C, 4.5% H, 2.4% N, 4.3% S and 12.3% O (by difference) on a dry, ash-free basis²⁵. Rietmeijer

and Mackinnon¹⁴ have suggested such macromolecular carbon as a possible source material that was metamorphosed to produce the graphite observed in the chondrites.

The facts may be summarized as follows. Graphite is more stable than either the poorly graphitized carbon or macromolecular, kerogen-like organic material observed in meteorites, and therefore should have survived to at least the same degree as did these other forms of carbon. Macromolecular interstellar carbon is found in carbonaceous chondrites, as are the somewhat more fragile 'grains' which contain trapped noble gas components indicative of their formation in various circumstellar environments. Graphite is not observed in the primitive meteorites except in trace amounts quite easily accounted for by the metamorphism of less stable carbonaceous precursors. Such metamorphic processes should not have destroyed pre-existing graphite grains. It therefore seems logical to conclude that graphitic carbon was not a major component of interstellar dust at the time the Solar System formed. Since graphite is not observed to form in R Corona Borealis stars, a circumstellar environment which should be more favourable for the production of graphite grains than a typical carbon-rich circumstellar outflow, we can also infer that newly formed carbon particles are not graphitic. For this reason we might be justified in concluding that graphite is likely to be only a minor component of the modern interstellar dust population. This conclusion may be consistent with a recent suggestion by Hecht²⁶ that the 220-nm bump arises from a population of extremely small graphite grains which would constitute only a small fraction of the total mass of the solid carbonaceous component of interstellar dust.

Most of the questions following from these rather straightforward interpretations can only be answered by detailed laboratory measurements of the appropriate materials and processes. For example, the optical properties of the macromolecular organic component of the meteorites and of QCCs need to be measured as a function of 'metamorphic grade' in order to determine the strength, width and stability of any potentially observable spectral features. Similarly, the effects of the space environment on such materials should be investigated; for example, the effects of irradiation by cosmic rays, X rays or ultraviolet light, the degree of reprocessing induced by shocks and the degree to which thermal annealing can cause 'graphitization'. In addition, it may be possible to use meteoritic data, such as measurements of the abundances of specific 'stellar' isotopic signatures, to place an upper limit on the efficiency with which grains are destroyed in the interstellar medium. Clearly the science of meteoritics can offer considerable data to astrophysicists. Closer cooperation and more communication between these communities is needed to ensure that the available data are used to their fullest potential.

This work was performed while I held a NASA NAS/NRC Research Management Associateship in the Solar System Exploration Division at NASA Headquarters. I acknowledge the helpful advice and suggestions of Bert Donn, Frans Rietmeijer and John Kerridge.

Received 22 May; accepted 4 September 1985.

- Hayatsu, R. & Anders, E. *Topics Curr. Chem.* **99**, 1–37 (1981).
- Robert, F. & Epstein, S. *Meteoritics* (abstr.) **15**, 355–356 (1980).
- Kerridge, J. *Earth planet. Sci. Lett.* **64**, 186–200 (1980).
- Srinivasan, B. & Anders, E. *Science* **201**, 51–56 (1978).
- Black, D. *Geochim. cosmochim. Acta* **36**, 377–394 (1972).
- Clayton, D. D. *Nature* **257**, 36–37 (1975).
- Clayton, D. D. *Q. Jl R. astr. Soc.* **23**, 174–212 (1982).
- Cameron, A. G. W. in *Interstellar Dust and Related Topics* (eds Greenberg, J. M. & van de Hulst, H. C.) 345–347 (Reidel, Boston, 1973).
- Black, D. & Pepin, R. *Earth planet. Sci. Lett.* **6**, 395–405 (1969).
- Junge, M. H. A. & Eberhardt, P. *Meteoritics* **14**, 439–441 (1979).
- Reynolds, J. H. & Turner, G. *J. geophys. Res.* **69**, 3263–3281 (1964).
- Clayton, D. D. & Ward, R. A. *Astrophys. J.* **193**, 397–399 (1974).
- Dodd, R. T. *Meteorites: A Petrochemical Synthesis* (Cambridge University Press, 1981).
- Rietmeijer, F. J. M. & Mackinnon, I. D. R. in *16th lun. planet. Sci. Conf. Abstr.* 700 (Lunar and Planetary Institute, Houston, 1985).
- Mathis, J., Rimpl, W. & Nordsieck, K. H. *Astrophys. J.* **217**, 425–433 (1977).
- Hong, S. S. & Greenberg, J. M. *Astr. Astrophys.* **88**, 194–202 (1980).
- Millar, T. J. & Duley, W. W. *Mon. Not. R. astr. Soc.* **183**, 177–185 (1978).
- Hoyle, F. & Wickramasinghe, N. C. *Astrophys. Space Sci.* **66**, 77–90 (1979).
- Sakata, A., Wada, S., Okutsu, Y., Shintani, H. & Nakata, Y. *Nature* **301**, 493–494 (1983).
- Hecht, J., Holm, A., Donn, B. & Wu, C. *Astrophys. J.* **280**, 228–234 (1984).

21. Stephens, J. R. *Astrophys. J.* **237**, 450-461 (1980).
22. Lewis, J. S. & Ney, E. P. *Astrophys. J.* **234**, 154-157 (1979).
23. Dwek, E. in *Proc. Workshop Interrelationships Among Circumstellar, Interstellar and Interplanetary Grains* (eds Nuth, J. & Stencl, R.) (NASA, Washington, DC, in the press).
24. Watson, W. D. *Rev. mod. Phys.* **48**, 513-552 (1976).
25. Hayatsu, R. *et al. Science* **207**, 1202-1204 (1980).
26. Hecht, J. H. *Astrophys. J.* (submitted).

Formation of mesosiderites by low-velocity impacts as a natural consequence of planet formation

John T. Wasson*† & Alan E. Rubin*

* Institute of Geophysics and Planetary Physics and † Departments of Earth and Space Sciences and of Chemistry and Biochemistry, University of California, Los Angeles, California 90024, USA

Mesosiderites are an enigmatic class of stony-iron meteorites composed of roughly equal amounts of pyroxenitic and basaltic silicates and metallic iron-nickel; olivine is rare. We suggest here that, during the period of planet formation, mesosiderites originated by the low-velocity collisions of large metallic core fragments with the surface of a differentiated asteroid-size body. Relative velocities of the asteroids native to this region were low ($\leq 1 \text{ km s}^{-1}$) because it was distant from massive protoplanets. However, a few differentiated asteroids were destroyed by high-velocity collisions with interlopers perturbed by protoplanets into highly eccentric orbits. These collisions reduced mantles and crusts to small silicate fragments but left cores in the form of large, durable metal fragments. Mesosiderite-like pyroxenite-basalt-metal mixtures were formed when large core fragments accreted at low velocities to the regolith of an intact asteroid; the metal contents of other regolithic regions remained low. Most of the olivine in mesosiderite and howardite breccias we attribute to the mantles of the disrupted parent bodies. The low ($12\text{--}20 \text{ mg g}^{-1}$) olivine contents indicate that the amount of debris accreted from disrupted asteroids was small relative to the volumes of the regoliths.

Because differentiated bodies with basalt-rich crusts are expected to have ultramafic mantles and metal-rich cores, models in which mesosiderite basalt and metal are proposed to have originated in the same body cannot easily explain the scarcity of olivine in these stony-iron meteorites. For example, one recent suggestion is that basaltic crustal blocks foundered and sank through a molten peridotite mantle and were then invaded by molten metal from the core¹. This model requires an implausible thermal history to explain the survival of basalt (melting temperature $\sim 1,400 \text{ K}$) exposed to molten mantle and core material ($T \geq 1,800 \text{ K}$).

The least recrystallized mesosiderites (Fig. 1a) are impact breccias, consisting of poorly sorted, angular, igneous silicate clasts, recrystallized breccia clasts and metal nodules in a matrix of metallic Fe-Ni and silicate grains²⁻⁵. The mesosiderites that have undergone the greatest thermal processing are clast-laden impact melt-rocks in which immiscible silicate and metal-sulphide liquids coexisted^{5,6}. Evidence for brief high-temperature events, inferred from mineralogical disequilibrium⁷, also support the proposal that mesosiderites are impact products.

The model invoking core accretion to basalt-rich regolith⁸ has not gained acceptance because relative impact velocities in the asteroid belt ($\sim 5 \text{ km s}^{-1}$) are too high to allow a projectile/target mass ratio near unity. Although previous workers⁹ recognized that low ratios would be expected at lower relative velocities ($\leq 1 \text{ km s}^{-1}$), no scenario leading to such low velocities has been proposed.

The differentiation of planetary bodies normally results in the extrusion of basaltic lavas. Pyroxenitic cumulates could lie immediately below these basalts. Spectral evidence of basalt on the asteroid 4 Vesta^{10,11} suggests that many asteroid-size bodies that formed within Vesta's orbit (2.4 AU from the Sun) were differentiated because most plausible asteroidal heat sources

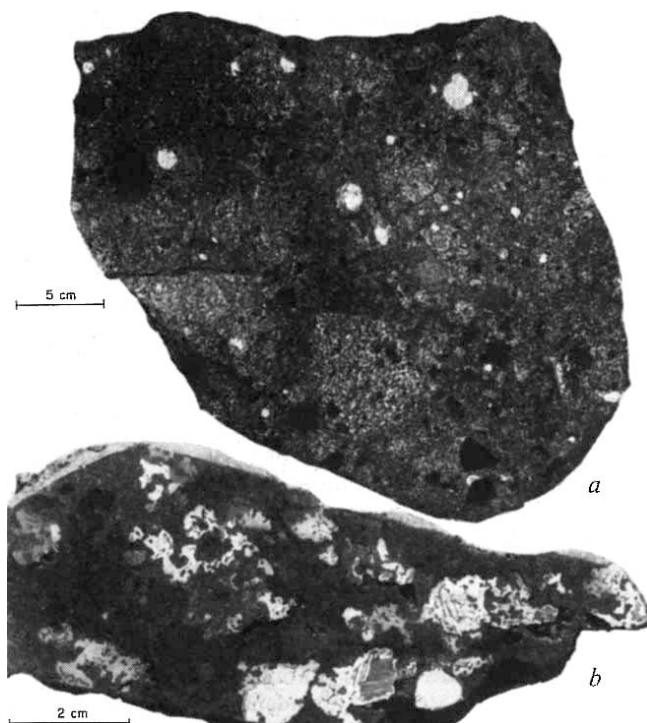


Fig. 1 a, Photograph of the Mt Padbury subgroup 1A mesosiderite (Smithsonian Institution specimen USNM 3454) showing angular silicate clasts, metal-rich breccia clasts and metal-sulphide nodules within a metal-silicate matrix. Metal grains have been finely comminuted and intimately mixed with silicates through a series of impacts. b, Photograph of the Estherville subgroup 3/4 A mesosiderite (Smithsonian Institution specimen USNM 1025) showing many large rounded metal nodules within a silicate matrix. Estherville suffered less extensive impact reworking than many other mesosiderites.

increase in efficiency with increasing proximity to the Sun. This is obvious for Sun-centred sources such as solar wind-induced electrical currents, but is also true for impact heating, because relative velocities are larger nearer the Sun. Heating of planetesimals by ^{26}Al is more efficient for objects forming earlier in the history of the Solar System; a simple model proposed by Safronov¹² suggests that planetesimals formed earlier nearer the Sun. Internal isochrons determined for eucrites and howardites¹³⁻¹⁶ show that the melting of these bodies occurred within 50 Myr of the formation of chondrites.

The oxygen-isotope compositions of the possibly martian SNC meteorites are only $0.4\% \delta^{17}\text{O}$ above the terrestrial fractionation line¹⁷, supporting the suggestion that primitive bodies that formed in the inner Solar System had O-isotope compositions near this line¹⁸. The O-isotope compositions of mesosiderites and related differentiated meteorites (eucrites, howardites and diogenites) lie equally near the terrestrial fractionation line¹⁹, suggesting that these meteorites also formed in the inner Solar System, perhaps inside 1.5 AU.

Wasson and Warren²⁰ cited evidence in support of the idea that the Moon formed by the circumterrestrial accretion of silicate fragments from differentiated asteroid-size bodies. It seems plausible that, at that time, more than 90% of the bodies in heliocentric orbits having radii near 1 AU were differentiated. We suggest that a similar fraction was differentiated at the mesosiderite formation location. Eucrites, howardites and mesosiderites are unlikely to be from Vesta²¹, because this asteroid is only one of a very large set of differentiated bodies. Fragmented members of this set are far more plausible parents.

According to the scenarios for planet formation proposed by Safronov¹² and Wetherill²², about 100 Myr was required between the stage of planetesimal formation by gravitational collapse to the growth of the inner planets to 95% of their present sizes.

At some epoch during planet formation, most bodies were in low-inclination, nearly circular orbits and average relative velocities in the inner Solar System were $\sim 1 \text{ km s}^{-1}$. Such low velocities would be preserved until 10^{26} -g, Mars-size bodies in eccentric orbits passed through the region. If some distant protoplanets were already Mars-size, some asteroid-size bodies in their vicinities could have been perturbed into eccentric orbits that caused them to pass through the mesosiderite formation region with relative velocities of $5\text{--}10 \text{ km s}^{-1}$. Impacts with these high-velocity interlopers fragmented a few differentiated native bodies into crust, mantle and core debris, the relative masses of these materials probably being about 10%, 65% and 25%, respectively. In those cases in which the high-velocity interloper was much smaller than the body it destroyed, the fragments of both bodies ended up in heliocentric orbits that were nearly circular. Weak crusts and mantles broke into small pieces; the more durable metal cores remained largely intact.

The accretion of a metallic projectile onto a basalt-rich regolith at a velocity of $\sim 1 \text{ km s}^{-1}$ should form large amounts of materials in which the metal/silicate ratio is approximately unity. Some blocks of metal or silicate would not be reduced to the millimetre size typical of mesosiderites, but those with dimensions $\geq 20 \text{ cm}$ might not be recognized as mesosiderites if they fell to Earth as individual meteorites. The textures of some mesosiderites attest to several generations of breccia formation (Fig. 1a); thus, a series of impacts may have reduced the mean grain sizes to those observed. Estherville (Fig. 1b) contains large (1–3 cm diameter) metal nodules and has probably experienced less extensive impact reworking than mesosiderites possessing fine-grained intimate mixtures of metal and silicate (for example, Mt Padbury, Emery, Mincy). Continuing impacts may have generated sufficient heat to cause metamorphism or even incipient melting in some mesosiderites.

Because, in a typical asteroid, the mass of the peridotitic mantle is expected to be about three times greater than the mass of metal in the core, we expect the proportion of accreted olivine to be more abundant than that of metal in the regolith. If olivine was mainly accreted as small debris, the amount of olivine in mesosiderites can be used as a measure of the fraction of the regolith composed of material derived from the accretion of one or more disrupted differentiated bodies. The median olivine content of mesosiderite silicates ($\sim 16 \text{ mg g}^{-1}$)²³ implies that $\sim 25 \text{ mg g}^{-1}$ of the regolith accreted from this source. The howardite regolith breccias provide a test of this scenario; they have chemical and O-isotope compositions nearly indistinguishable from those of the silicate portions of the mesosiderites, and probably formed either on the mesosiderite body or on a neighbouring body. Their median olivine content ($\sim 12 \text{ mg g}^{-1}$; ref. 24) is similar to that of mesosiderites.

Large ($\sim 1 \text{ cm}$) olivine clasts in Mt Padbury, Mincy, Pinnaroo, Vaca Muerta, and West Point have compositions ranging from Fa_{8-12} (refs 2, 25–27); in Emery such clasts range from Fa_{18-28} (ref. 26). The compositional range of the large clasts is narrower than that of all mesosiderite olivine clasts (Fa_{9-42})²⁸ and is similar to that (Fa_{10-30}) calculated for the lunar mantle from magma ocean models²⁹. These observations are consistent with a mantle origin for mesosiderite olivine. Olivine composition is best preserved in large clasts; fine-grained olivine tended to become more ferroan as a result of equilibration with the much more abundant indigenous mafic silicates in the regolith.

If a negligible ($<10\%$) fraction of mesosiderite olivine is indigenous, the observed 16 mg g^{-1} of mantle olivine implies that $\sim 5 \text{ mg g}^{-1}$ of the regolith consisted of core materials. This low abundance is not inconsistent with the fact that metal accounts for $\sim 500 \text{ mg g}^{-1}$ of mesosiderites if we assume that the metal from disrupted asteroids primarily resided in large ($>100 \text{ m}$) fragments. Mesosiderites formed where these objects accreted, whereas metal contents of other regolithic regions remained low, as in howardites. Because the metal content of the mesosiderites is 100 times greater than that estimated for the entire regolith, mesosideritic regolith is inferred to have formed only 1% of the parent body surface.

There are two compositional subgroups (A and B) of mesosiderites, possessing different modal abundances of orthopyroxene, plagioclase and tridymite^{5,23,30}. The somewhat different mean Ir/Ni ratios ($\times 10^5$) of the metal in these groups (3.6 ± 0.91 and 5.1 ± 1.2 , respectively)⁸ suggest that metal core fragments accreted to at least two distinct regolithic regions.

Mesosiderite metal compositions are roughly chondritic^{8,31}. Most cores seem to have formed by fractional crystallization³², but most core fragments are approximately chondritic. The Rayleigh equation can be used to show that nearly half the volume of a metal core would have siderophile ratios within a factor of 3 of the chondritic ratio. This is the case even for elements having extreme solid-liquid distribution coefficients (k) such as Au ($k=0.4$) and Ir ($k=3$).

The moderately high abundance for mesosiderites among meteorite falls can be accounted for by biases resulting from parent body ejection, space erosion and the rigors of atmospheric passage, all of which favour the tough, metal-rich mesosiderites over the more friable howardites.

The only mesosiderite property seemingly incompatible with a near-surface origin is the apparently slow cooling rate below $\sim 800 \text{ K}$ inferred from metallographic evidence. Powell³ and Hewins³³ estimated this rate to be $\leq 1 \text{ K Myr}^{-1}$. Nevertheless, it is possible that it was not monotonic cooling, but low-temperature ($600\text{--}700 \text{ K}$) annealing or thermal cycling for long time periods that produced the metallographic structures. Even if slow cooling actually occurred, the burial depth^{34,35} could be as low as $\sim 15 \text{ km}$ if the thermal diffusivity was similar to that of the lunar regolith ($10^{-4} \text{ cm}^2 \text{ s}^{-1}$)³⁶. Such a regolith depth is plausible for a body accreting material at low relative velocities.

Thus, it seems probable that mesosiderites formed when asteroid-size bodies were still growing. We suggest that many such bodies in the inner Solar System differentiated before accumulation was complete. If a few of these differentiated bodies were disrupted by high-velocity interlopers, it is inevitable that some of their fragments would accrete at low velocities to the basalt-rich regoliths of bodies that avoided disruption. Olivine contents reflect the fraction of the regolith from this source. Mesosiderite-like materials were produced in those rare regolith locations where large core fragments accumulated. Selection biases enhanced the flux of mesosiderite-like breccias relative to that of howardite-like breccias onto the Earth's surface.

We thank G. W. Wetherill, R. H. Hewins and J. S. Delaney for useful comments on earlier versions of this manuscript and T. B. Thomas for the photographs of Mt Padbury and Estherville. This work was supported by NASA grant NAG 9-40.

Received 9 April; accepted 26 September 1985.

- Greenberg, R. & Chapman, C. R. *Icarus* **57**, 267–279 (1984).
- McCall, G. J. H. *Miner. Mag.* **35**, 1029–1060 (1966).
- Powell, B. N. *Geochim. cosmochim. Acta* **33**, 789–810 (1969).
- Powell, B. N. *Geochim. cosmochim. Acta* **35**, 5–34 (1971).
- Floran, R. J. *Proc. lunar planet. Sci. Conf.* **9**, 1053–1081 (1978).
- Floran, R. J., Caulfield, J. B. D., Harlow, G. E. & Prinz, M. *Proc. lunar planet. Sci. Conf.* **9**, 1083–1114 (1978).
- Delaney, J. S., Nehru, C. E., Prinz, M. & Harlow, G. E. *Proc. lunar planet. Sci. Conf.* **12B**, 1315–1342 (1981).
- Wasson, J. T., Schaudy, R., Bild, R. W. & Chou, C.-L. *Geochim. cosmochim. Acta* **38**, 135–149 (1974).
- Hewins, R. H. *Workshop on Lunar Breccias and Soils and Their Meteoritic Analogs* (eds Taylor, G. J. & Wilkening, L. L.) 49–53 (LPI Tech. Rep. 82-02, Lunar Planetary Institute, Houston, 1982).
- McCord, T. B., Adams, J. B. & Johnson, T. V. *Science* **168**, 1445–1447 (1970).
- Larson, H. P. & Fink, U. *Icarus* **26**, 420–424 (1975).
- Safronov, V. S. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets* (Nauka, Moscow; transl. Israel Program. Sci. Transl. NASA TTF-677, 1969).
- Allègre, C. J., Birck, J. L., Fourcade, S. & Semet, M. *Science* **187**, 436–438 (1975).
- Birck, J. L. & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 37–51 (1978).
- Papanastassiou, D. A. & Wasserburg, G. J. *Lunar Sci.* **7**, 665–667 (1976).
- Papanastassiou, D. A., Rajan, R. S., Huneke, J. C. & Wasserburg, G. J. *Lunar Sci.* **5**, 583–585 (1974).
- Clayton, R. N. & Mayeda, T. K. *Earth planet. Sci. Lett.* **62**, 1–6 (1983).
- Wasson, J. T. & Wetherill, G. W. in *Asteroids* (ed. Gehrels, T.) 926–974 (University of Arizona Press, Tucson, 1979).
- Clayton, R. N., Onuma, N. & Mayeda, T. K. *Earth planet. Sci. Lett.* **30**, 10–18 (1976).
- Wasson, J. T. & Warren, P. H. *Lunar planet. Sci.* **10**, 1310–1312 (1979).
- Consolmagno, G. J. & Drake, M. J. *Geochim. cosmochim. Acta* **41**, 1271–1282 (1977).
- Wetherill, G. W. *Astr. Astrophys.* **18**, 77–113 (1980).

23. Prinz, M., Nehru, C. E., Delaney, J. S., Harlow, G. E. & Bedell, R. L. *Proc. Lunar Planet. Sci. Conf.* 11, 1055-1071 (1980).
24. Delaney, J. S., Prinz, M. & Takeda, H. *Proc. lunar planet. Sci. Conf.* 15, C251-C288 (1984).
25. Hewins, R. H. & Ulmer, G. C. *Geochim. cosmochim. Acta* 48, 1555-1560 (1984).
26. Mittlefehldt, D. W. *Earth planet. Sci. Lett.* 51, 29-40 (1980).
27. Prior, G. T. *Miner. Mag.* 18, 151-172 (1918).
28. Delaney, J. S., Nehru, C. E. & Prinz, M. *Proc. lunar planet. Sci. Conf.* 11, 1073-1087 (1980).
29. Warren, P. H. & Wasson, J. T. *Proc. lunar planet. Sci. Conf.* 10, 2051-2083 (1979).
30. Hewins, R. H. *Proc. lunar planet. Sci. Conf.* 15, C289-C297 (1984).
31. Bild, R. W., Robinson, K. L., Scott, E. R. D. & Prinz, M. *Meteoritics* 18, 266-277 (1983).
32. Scott, E. R. D. *Geochim. cosmochim. Acta* 36, 1205-1236 (1972).
33. Hewins, R. H. *Proc. lunar planet. Sci. Conf.* 14, B257-B266 (1983).
34. Taylor, G. J. *et al. Geochim. cosmochim. Acta* 43, 323-337 (1979).
35. Rubin, A. E. *et al. Geochim. cosmochim. Acta* 45, 2213-2228 (1981).
36. Langseth, M. G., Keihm, S. J. & Peters, K. *Proc. lunar. Sci. Conf.* 7, 3143-3171 (1976).

A large-scale meridional circulation in the convective zone

E. Ribes, P. Mein & A. Mangeney

Observatoire de Paris, Section d'astrophysique de Meudon,
92190—Meudon, France

Many attempts have been made to detect, at the solar photospheric level, large-scale motions related to the Sun's global convection and dynamo (see ref. 1). Because these motions are likely to be influenced by solar rotation, one might expect that they would take the form of 'bananas' aligned with the solar rotation axis², as a result of the Taylor-Proudman constraint. The Paris observatory has one of the longest series of spectroheliograms, dating back to 1919. In view of the current interest in the history of solar activity, we have started to digitize the collection³. This allows us to measure rotational rate and meridional drift for individual sunspots with an accuracy of a few ms^{-1} for the best seeing conditions at Meudon. A new phenomenon, which may lead us drastically to revise our ideas on the large-scale convection of the Sun, can immediately be seen. Newly-born sunspots trace a roughly axisymmetric meridional circulation, in the form of four zonal bands, with relatively large velocity amplitude. If these drifts are assumed to trace fluid motions at some level in the solar envelope, the resulting circulation pattern cannot be associated with the 'banana' cells mentioned above. We report here results based on two sets of data; one involves a detailed analysis of two periods, each covering five solar rotations while the other is essentially qualitative and results from eye-estimates of the spot meridional drifts over much longer time sequences.

The first period (Carrington rotations 1547 to 1551) spans sunspots at the maximum of the solar cycle 20th, the second (Carrington rotations 1587 to 1591) corresponds to the declining phase. Twenty new bipolar groups were born on the visible disk for each period. A third sequence covering the summer of 1967 (ascending phase) has been partially analysed and confirms the following results.

A meridional circulation pattern can be seen in Fig. 1. This is due to young sunspots, and it is larger than in any possible experimental error. It consists of four latitude bands, two in each hemisphere, each pair with opposite directions of circulation. Furthermore, the sense of the meridional drift within each band reverses during the solar cycle. For example, in 1969, sunspots born between latitudes 10° and -21° move towards the equator while sunspots born beyond these latitudes move towards the poles. In 1972, sunspots born within $14^\circ 5'$ and -17° move towards the poles while outside this range, they move towards the equator. Meridional velocities range from zero to 100 m s^{-1} (1° day^{-1} corresponds to 130 m s^{-1}). This value is compatible with convective velocities predicted for giant cells⁴.

One property of the meridional circulation is that it is roughly axisymmetric. Although sunspots are often found at certain preferred longitudes, our data sample a fairly wide range of longitude, especially in the declining phase of the cycle; the observed drifts do not seem to depend on longitude. However, this property of axisymmetry should be confirmed over a larger statistical sample.

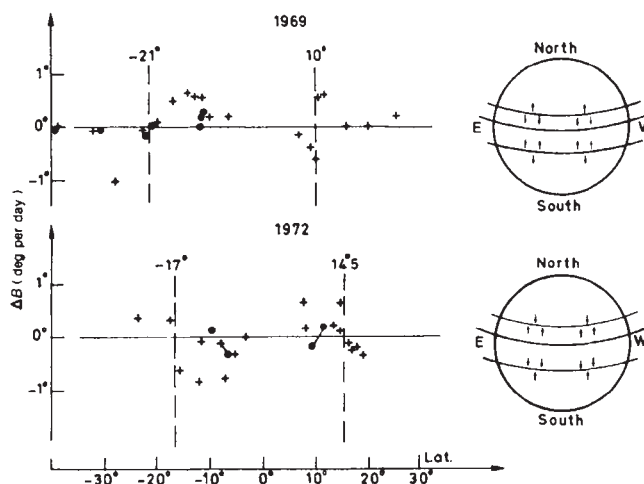


Fig. 1 The meridional circulation traced by young sunspots plotted as a function of latitude for the 1969 and 1972 sequences. $\times$, Leading sunspots; $\bullet$, following sunspots. In the northern hemisphere, a positive meridional drift ($\Delta B > 0$) indicates that spots are moving towards the north pole, while a negative drift ($\Delta B < 0$) corresponds to a displacement of spots towards the equator.

The rotational rate of young spots is nearly constant with latitude. The synodic leader's rate ($14^\circ 22'$ per day in 1969) is larger than the follower's, as expected from the rise of the magnetic loop (the difference is about $1^\circ 5'$ for the first day and drops rapidly during the next days). The mean deviation, σ , of $0^\circ 7'$ for the leaders and $0^\circ 5'$ for the followers is mainly due to some low latitude spots (at the boundary of the two latitude bands) and does not reflect the differential rotation. Note that the large σ value is partly due to the difficulty of observing young spots on the first day of their emergence. The rotation rate, Ω , at the depth where the spots are anchored should be determined from the centre of gravity of the emerging magnetic loop. But because this is unknown, we used the midpoint between the leaders and the followers. Ω is then approximately equal to the surface equatorial rate. This result is consistent with the deep solid rotation implied by the observed rotational splitting of global oscillations^{5,6}. It also indicates that young spots are better tracers of the deep convective layers than older spots which exhibit a pronounced differential rotation^{7,8}.

The average rotational rate increases by 5% between 1969 and 1972, an increase comparable with the σ value of young spot rotation. This effect is in qualitative agreement with the report⁹ that the solar rotation is faster at the minimum of the cycle. Our sample is too small to allow a quantitative comparison to be made, however.

The question then arises, of whether this meridional circulation of the young spots traces fluid motions at some level in the solar envelope. There are several arguments which suggest that it does.

First, sunspots usually emerge at the surface as simple dipolar magnetic features. When the magnetic loop rises, the leading and the following spots diverge and, a few days later, the following spot starts to fragment and gradually disperses into the general surrounding magnetic field, leaving only the leading spot as an identifiable feature. If one accepts that sunspots are anchored deeply in the convective zone¹⁰, then the drag exerted by the surrounding plasma on the magnetic tubes should be much stronger when the magnetic structure is more complex, that is in the later stage of the spot's life. Thus, newly-born spots are probably the best tracers of fluid motions in the deep convection zone where the dynamo process is supposedly acting.

Second, the spot circulation seems to be related to the large-scale, long-lived, magnetic structures outlined by the *H α* filaments. The procedure used to look at these structures is

Fig. 2 Evidence of azimuthal rolls (dashed arrows) from the observed meridional circulation of young sunspots (solid arrows) and the drift of long-lived $H\alpha$ filament trajectories (solid lines). The direction of rotation seems to be associated with the magnetic polarity. A new roll appears every 2 or 3 yr. Sunspot activity seems to be concentrated near the updrafts (u) more than near the downdrafts (d).

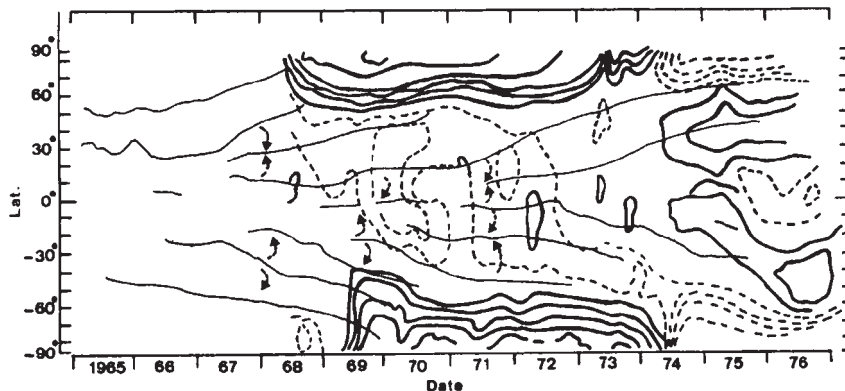
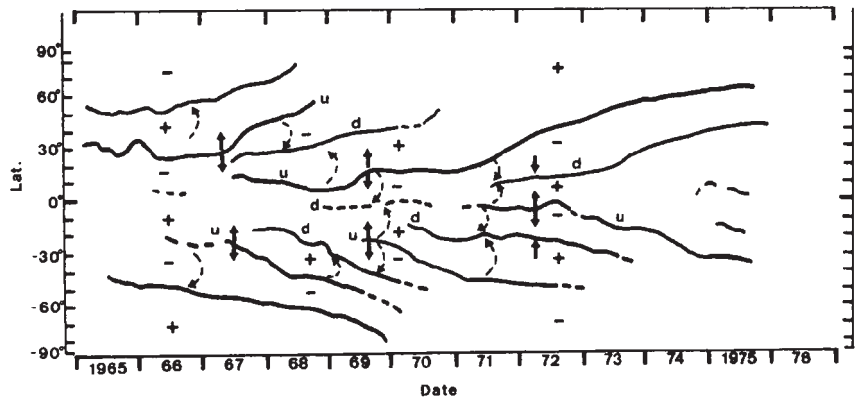


Fig. 3 Contour maps of differences from average differential rotation (intervals are ± 16 , ± 8 and $\pm 4 \text{ m s}^{-1}$) obtained by Snodgrass and Howard¹². Dashed lines and solid lines correspond to smaller and larger rotational shear. The azimuthal rolls (thin lines) seem to influence the angular momentum transport by reducing the rotational shear. The direction of rotation of the azimuthal rolls is indicated by the dashed arrows.

similar to that used by Makarov¹¹. A mean latitude is assigned to every filament band present on the solar surface during one rotation. A three-rotation-average for each band is then taken. The filaments bands drift with time. If we accept MacIntosh's claim that $H\alpha$ filaments outline the boundaries of regions of unipolar magnetic fields, we can determine an azimuthally-averaged magnetic pattern of alternating polarity. We have disregarded short-lived $H\alpha$ filaments which introduce noise in the data (in contrast with Makarov). As shown in Fig. 2 (see also ref. 11), the magnetic pattern is not random and consists of large-scale azimuthal magnetic regions. Moreover, the magnetic pattern drawn by $H\alpha$ filaments (solid lines in Fig. 2) fits the critical latitudes of the zonal meridional circulation. It is difficult to see how such a pattern could be established other than by large-scale motions. The remarkable coincidence between the circulation pattern and the magnetic pattern suggests that we are looking at azimuthal rolls.

Third, the circulation pattern reflected by young spots interacts with the transport of angular momentum. Snodgrass and Howard¹² have noticed periodic changes in the differential rotation close to sunspot maximum. By superimposing the azimuthal pattern on their data (Fig. 3), it can be seen that the presence of rolls is restricted to the latitudes which have less rotational shear. This is a natural consequence of a meridional circulation in the form of several counter rotating cells.

From these converging arguments, a new picture of the dynamics of the convective zone and the dynamo action emerges. Four to six azimuthal rolls are present, having widths of $\sim 20^\circ$ – 40° with a lifetime of several years. The direction of rotation of a roll is associated with its magnetic polarity. Every 2 or 3 yr, a roll dislocates and a new one appears. The appearance of rolls and their dislocations seems to be related to sunspot activity. In particular, new spots are found mostly in the vicinity of the boundaries between two adjacent rolls. They arise preferentially where the two rolls are diverging rather than converging. This suggests that magnetic fields are expelled by upward motions¹³. After each bifurcation of a roll (in 1967, 1969, 1970 and 1971), a new one appears and the updraft situation moves closer to

the equator. This could explain the shift in latitude observed in the so-called Maunder butterfly diagram of sunspots.

Is there any connection with the giant convective cells which have been invoked³ but never detected? It is hard to give a definitive answer as our method of analysis is biased towards larger-scale motions and precludes the detection of smaller-scale motions such as the giant cells mentioned previously. However, the velocities we have observed are comparable in magnitude to convective velocities calculated for the deep convection zone⁴. This is a very strong argument in favour of a convective origin for the meridional circulation of young spots. If this is the case, we must relax the Proudman–Taylor's constraint. The radial rotational shear, inferred from helioseismology⁶, is too weak in the convective zone to be relevant. On the other hand, a toroidal magnetic field of 10^3 G could force the large-scale convection to take the form of rolls aligned with the azimuthal direction¹⁴.

We thank Drs J. M. Robillot, B. Leroy and P. A. Gilman for stimulating discussions, also Dr C. H. Barrow for checking our English. $H\alpha$ synoptic maps have been edited by Observatoire de Paris, under the supervision of Dr M. J. Martres. The spectroheliographic plates have been digitized with the PDS microdensitometer (Institut d'optique, Orsay) and analysed using a Vax computer (Paris Observatory). This work was supported by an INAG grant.

Received 14 June; accepted 3 September 1985.

1. Schröter, E. H. *Sol. Phys.* **100** (in the press).
2. Gilman, P. A. in *The Sun as a Star* (ed. Jordan, St) 231–252 (CNRS, France and NASA, Washington DC, 1981).
3. Mein, P. & Ribes, E. *Astr. and Astrophys.* (submitted).
4. Weiss, N. O. *MON. Not. R. astr. Soc.* **128**, 227 (1964).
5. Duvall, T. L. Jr & Harvey, J. W. *Nature* **310**, 19–22 (1984).
6. Duvall, T. L. Jr *et al.* *Nature* **310**, 22–25 (1984).
7. Balthazar, H. & Wohl, H. *Sol. Phys.* **88**, 71–75 (1983).
8. Howard, R., Gilman, P. A. & Gilman, P. I. *Astrophys. J.* **283**, 373–384 (1984).
9. Gilman, P. A. & Howard, R. *Astrophys. J.* **283**, 385–391 (1984).
10. Parker, E. N. in *Cosmical magnetic fields* (eds Marshall, W. & Wilkinson, D. H.) 165 (Oxford University Press, 1979).
11. Makarov, V. I. *Sol. Phys.* **93**, 393–396 (1984).
12. Snodgrass, H. B. & Howard, R. *Sol. Phys.* **95**, 221–228 (1985).
13. MacIntosh, P. S. & Wilson, P. R. *Sol. Phys.* **97**, 59–80 (1985).
14. Gough, D. in *Problems of Stellar Convection* (eds Spiegel, E. A. & Zahn, J. P.) 71 (Lecture Note in Phys., Springer, Berlin, 1976).

Parity-violating energy differences of chiral minerals and the origin of biomolecular homochirality

G. E. Tranter

Theoretical Chemistry Department, University of Oxford, Oxford OX1 3TG, UK

The biochemistry of terrestrial organisms is based on chiral (handed) molecules, with one of the two possible series of enantiomers (mirror-image isomers) being predominant. Specifically, terrestrial biochemistry is homochirally supported by L- α -amino acids and D-sugars to the almost complete exclusion of the enantiomeric D- α -amino acids and L-sugars. This particular homochiral selection may be a result of very small differences between the electronic energies of enantiomeric prebiotic molecules due to the parity-violating weak interactions. The energy differences are, however, so small for simple chiral molecules, that the propagation of homochirality would require a dissymmetry amplification mechanism involving both large quantities of reactants and a long reaction time. An alternative theory, presented here, considering the effects of the parity-violating weak interactions in crystalline enantio-selective prebiotic catalysts, such as the clay silicates, may require considerably less amplification.

The natural competitive selection of a homochiral biochemistry is accounted for generally by elaborations of the stereochemical 'lock-and-key' hypothesis of Fischer¹, but conventionally the terrestrial selection of the L- α -amino acid and D-sugar series over the classically equally viable enantiomeric alternative has been ascribed to chance. The efficient formation of prototype biomolecular polymers during the prebiotic-biotic transition requires the existence of a chemistry approaching homochirality²⁻⁴, and, until recently, the mechanisms proposed for its evolution from a prebiotic racemic geochemistry were equivocal, either of the two enantiomeric homochiral chemistries being selected with equal probability⁵. The violation of parity by the weak interactions provides an alternative, globally consistent, basis for the systematic and determinate selection of the observed homochirality without recourse to chance events.

Through the connection established with electromagnetism^{6,7} the parity-violating weak interactions cause a chiral molecular system and its enantiomer to have slightly inequivalent energies⁸⁻¹¹. Effectively there is a small parity-violating energy difference (PVED), ΔE_{pv} , between the corresponding state energies of a chiral system and its enantiomeric system, which preferentially stabilizes one with respect to the other. The PVED may occur for any chiral system, including transient forms, and can, therefore, manifest itself as an imbalance in the quantities of enantiomeric species in classically racemic equilibrium or as an inequivalence in the rate constants of enantiomeric reactions.

The intrinsic PVED between enantiomeric species and reactions involved in the prebiotic-biotic transition may have specifically selected the known terrestrial homochirality. Recent calculations indicate that the natural L-enantiomers of a series of α -amino acids are stabilized relative to their corresponding unnatural D-enantiomers for the conformations preferred in aqueous media¹²⁻¹⁵. A similar stabilization of the natural enantiomers has been determined for a set of polypeptide/protein fragments in both the α -helix and β -sheet conformations^{12,13} and also for a possible prebiotic reaction to form a precursor of L-alanine¹⁶.

The calculated values of PVED for small chiral molecules are exceedingly small, typically $\sim 10^{-20}$ atomic units (a.u., equivalent to 'hartrees'), although other conceivable systematic chiral influences are generally even weaker. Nonetheless, several kinetic mechanisms that amplify dissymmetry to yield homochirality have been suggested¹⁷⁻²⁰, and it has been determined that even a PVED as small as $\sim 10^{-20}$ a.u. may be sufficient as a selector. However, latest estimates²⁰ of the reaction conditions

required for such an amplification are a lake of volume 4×10^9 litres ($1 \text{ km} \times 1 \text{ km} \times 4 \text{ m}$) which remains homogeneous for some 10,000 yr. While such conditions are possible in the prebiotic era, they must be regarded as exceptional.

The essential difficulty is the minuteness of the intrinsic PVED for small prebiotic chiral molecules. Giant chiral molecules such as polymers, and especially crystal lattices, may have much larger PVED, and may exhibit the resulting dissymmetries at a detectable level without the intervention of amplification mechanisms. For example, the ratio of the numbers of left (*l*) and right (*d*)-handed enantiomorphous crystals (*l/d*), each with *n* unit cells, under chemical equilibrium at temperature *T* is given by

$$l/d = \exp(-\Delta E_{pv}^{(n)}/kT) \quad (1)$$

where *k* is the Boltzmann constant, and the PVED, $\Delta E_{pv}^{(n)}$, is the energy of a (*l*)-enantiomer crystal relative to a (*d*)-enantiomer crystal of the same size (*n*). Taking the first-order approximation^{21,22} that the overall PVED, $\Delta E_{pv}^{(n)}$, has a linear relationship with the number of unit cells, *n*, it can be expressed in terms of the mean PVED per unit, ΔE_{pv}^I , or the advantage factor $\varepsilon = -\Delta E_{pv}^I/kT$

$$\Delta E_{pv}^{(n)} = n\Delta E_{pv}^I = -nekT \quad (2)$$

whereupon the dissymmetry ratio becomes

$$l/d = e^{n\varepsilon} \approx 1 + n\varepsilon \quad |n\varepsilon| \ll 1 \quad (3)$$

A corresponding result can be obtained from a kinetic approach using the progressive accumulation model²¹, the PVED then referring to the transition states for crystal growth. The enantiomeric excess, Ω_{crys} , a more convenient measure of the dissymmetry, is related by

$$\begin{aligned} \Omega_{\text{crys}} &= \frac{l-d}{l+d} \\ &= \frac{e^{n\varepsilon} - 1}{e^{n\varepsilon} + 1} \\ &\approx \frac{n\varepsilon}{2} \quad |n\varepsilon| \ll 1 \end{aligned} \quad (4)$$

For the large number of units, *n*, possible in crystal lattices, even very small non-zero PVED per unit can be enhanced to give a detectable difference in the quantities of left- and right-handed forms of a chiral crystal. Remarkably, such a difference has been documented for naturally occurring quartz²³ where, over a total collection of 16,807 crystals, a 1% enantiomeric excess of *l*(-)-quartz has been found at all the locations sampled in the Americas, Europe and Asia, with a 90% statistical confidence level that the bias was not a matter of chance. Using a value of the advantage factor $\varepsilon \approx 10^{-17}$, appropriate for a unit PVED, ΔE_{pv}^I , of a small chiral structure composed of light atoms ($\sim 10^{-20}$ a.u.), the 1% enantiomeric excess of terrestrial quartz requires $n \approx 10^{15}$ in equation (4). A quartz crystal containing 10^{15} unit cells has the dimensions $\sim 0.1 \text{ mm}$ along each main edge, and is, therefore, very easily attainable. Consequently, the observed natural quartz bias may be evidence of the existence of PVED effects in the physical world, although, at present, it is impractical to perform the necessary calculations of PVED for quartz to confirm this.

An alternative explanation, that the excess of *l*(-)-quartz was laid down in the post-biotic period and was influenced by the already existing biomolecular homochirality is unlikely. The elevated temperatures associated with the hydrothermal crystallization of quartz are unfavourable, the α -amino acids, for example, racemizing at $\sim 160^\circ \text{C}$ in aqueous solution with a half-life of only a few hours²⁴. Some experimental evidence has also been reported²⁵ on the transference of β -ray dissymmetry to the crystallization of the chiral salt sodium ammonium tartrate

under radiolysis, although its implications for the crystallization of quartz and other minerals are unknown.

Quartz is only one of the many chiral crystalline silicates to be found as minerals of the Earth's surface, and each may exhibit, to differing degrees, analogous enantiomeric excesses. The role of terrestrial minerals, especially the layer-structured silicate clays, in the prebiotic-biotic transition has long been discussed^{26,27}, and interest has recently been furthered by the discovery that clays can absorb mechanical and electromagnetic energy and release it slowly over periods of days²⁸. The most obvious method by which minerals can influence the prebiotic-biotic transition is through their action as catalysts, however to produce an enantiomeric excess in chiral product molecules from racemic/achiral reactants, it is necessary not only to have an enantiomeric excess in the chiral catalyst, but also that it must act enantio-selectively. This enantioselectivity is common, and is illustrated by the preferential adsorption of L-alanine from a racemic mixture by L-(+)-quartz^{29,30}.

For a simple catalytic reaction transforming racemic/achiral reactants into the L- and D-forms of a chiral product with rate constants k^L and k^D respectively, the degree of enantio-selectivity of a particular catalyst can be represented by its enantio-selectivity factor, S

$$S = \frac{k^L - k^D}{k^L + k^D} \quad (5)$$

The enantio-selectivity factors for the left-handed, S_L , and right-handed, S_D , forms of otherwise equivalent chiral catalysts will be different but related by $S_L = -S_D$.

The combination of an enantiomeric excess, Ω_{crys} , of a chiral catalyst (composed of crystals of equal size and surface area), with an enantio-selectivity $S = S_L$ ensures that for such a reaction there is an overall enantiomeric excess, Ω_{prod} , in the numbers of left(L)- and right(D)-handed product molecules formed from a racemic/achiral reactant mixture which is given by

$$\Omega_{\text{prod}} = \frac{L - D}{L + D} = S\Omega_{\text{crys}} \quad (6)$$

If the catalyst is either racemic ($\Omega_{\text{crys}} = 0$) or not enantio-selective ($S = 0$), there is no transference of any chiral bias.

Taking $\Omega_{\text{crys}} = 0.01$, in keeping with the observed 1% enantiomeric excess of terrestrial quartz, and a conservative 0.1% enantio-selectivity ($S = 10^{-3}$), the resulting enantiomeric excess in the chiral product molecules will amount to some 0.001% ($\Omega_{\text{prod}} = 10^{-5}$); this is considerably larger than the corresponding enantiomeric excess due to the inherent PVED of the product molecules themselves, which typically only amounts to $|\Omega_{\text{prod}}| \approx 10^{-17}$ at ambient temperatures.

Whilst the chemistry resulting from chirally biased catalysis is not totally homochiral ($|\Omega_{\text{prod}}| = 1$) it is a great improvement over alternative models, and would require only minor subsequent dissymmetry amplification, which could be achieved by the various proposed mechanisms¹⁷⁻²⁰ using smaller reaction volumes and less time than that required by comparable models. Alternatively, chirally biased catalysis by minerals could form an integral part of a dissymmetry amplification mechanism, further ensuring the rapid propagation of homochirality. Enantio-selective frameworks in possible dissymmetry amplification processes have been experimentally demonstrated by the chirally biased occlusion of α -amino acids into glycine crystals³¹ and also in template-directed oligomerization of biomolecules².

Throughout the above discussion it has been implicitly assumed that the enantiomeric excess and selectivity of the catalyst are such that the bias is towards the observed natural enantiomers of the principal biomolecules rather than their unnatural enantiomeric counterparts. Unfortunately, at present, the validity of this assumption cannot be determined because there is little detailed knowledge of the chemical processes of the prebiotic-biotic transition, and thus the specific relevant catalysts and reactions cannot be identified. Nonetheless, the model's ability to produce relatively large enantiomeric excesses

with ease warrants careful consideration and further investigation.

In conclusion, the indirect globally systematic influence of the weak interactions on the prebiotic-biotic transition, through the parity-violating energy differences in crystalline enantio-selective catalysts, may have specifically selected the observed terrestrial biomolecular homochirality. The role of clays in the prebiotic-biotic transition is still under debate, but their inherent ability to ensure the emergence of a homochiral biochemistry is further evidence in their favour.

Received 27 June; accepted 6 September 1985.

1. Fischer, E. *Chem. Ber.* **27**, 2985-2993, 3189-3232 (1894).
2. Joyce, G. F. et al. *Nature* **310**, 602-604 (1984).
3. Idelson, M. & Blout, E. R. *J. Am. chem. Soc.* **80**, 2387-2393 (1958).
4. Blair, N. E., Dirbas, F. M. & Bonner, W. A. *Tetrahedron* **37**, 27-29 (1981).
5. Mason, S. F. *Nature* **311**, 19-23 (1984).
6. Weinberg, S. *Phys. Rev. Lett.* **19**, 1264-1266 (1967).
7. Salam, A. in *Proc. 8th Nobel Symp. Elementary Particle Physics* (ed. Svartholm, N.) 367 (Almqvist and Wiksell, Stockholm, 1968).
8. Rein, D. W. *J. molec. Evol.* **4**, 15-22 (1974).
9. Zel'dovich, B. Ya., Saakyan, D. B. & Sobel'man, I. I. *Soviet Phys. J. exp. theor. Phys.* **25**, 94-97 (1977).
10. Rein, D. W., Hegstrom, R. A. & Sandars, P. G. H. *Phys. Lett.* **A71**, 499-502 (1979).
11. Hegstrom, R. A., Rein, D. W. & Sandars, P. G. H. *J. chem. Phys.* **73**, 2329-2341 (1980).
12. Mason, S. F. & Tranter, G. E. *Molec. Phys.* **53**, 1091-1111 (1984).
13. Mason, S. F. & Tranter, G. E. *Proc. R. Soc. A* **397**, 45-65 (1985).
14. Tranter, G. E. *Chem. Phys. Lett.* **120**, 93-96 (1985).
15. Tranter, G. E. *Molec. Phys.* (in the press).
16. Tranter, G. E. *Chem. Phys. Lett.* **115**, 286-290 (1985).
17. Frank, F. C. *Biochim biophys. Acta* **11**, 459-463 (1953).
18. Kondepudi, D. K. & Nelson, G. W. *Phys. Rev. Lett.* **50**, 1023-1026 (1983).
19. Kondepudi, D. K. & Nelson, G. W. *Phys. Lett.* **106A**, 203-206 (1984).
20. Kondepudi, D. K. & Nelson, G. W. *Nature* **314**, 438-441 (1985).
21. Yagumata, Y. *J. theor. Biol.* **11**, 495-498 (1966).
22. Keszthelyi, L. *Phys. Lett.* **64A**, 287-288 (1977).
23. Palache, C., Berman, H. & Frondel, C. in *Dana's System of Mineralogy*, 7th edn, Vol. 3, 16 (Wiley, New York, 1962).
24. Thiemann, W. & Darge, W. *Origins Life* **5**, 263-283 (1974).
25. Kovacs, K. L. *Origins Life* **9**, 219-233 (1979); **11**, 37-52 (1981).
26. Cairns-Smith, A. G. in *Genetic Takeover and the Mineral Origins of Life* (Cambridge University Press, 1982).
27. Schwartz, A. W. & Orgel, L. E. *J. molec. Evol.* **21**, 299-300 (1985).
28. Coyne, L., Sweeney, M. & Hovatter, W. *J. Lumin.* **28**, 395-409 (1983).
29. Kavasmaneck, P. R. & Bonner, W. A. *J. Am. chem. Soc.* **99**, 44-50 (1977).
30. Furuyama, S., Sawada, M., Machiya, K. & Morimoto, T. *Bull. chem. Soc. Jap.* **55**, 3394-3397 (1982).
31. Weissbuch, I. et al. *Nature* **310**, 161-162 (1984).

New higher primates from the early Miocene of Buluk, Kenya

R. E. F. Leakey* & A. Walkert†

* National Museums of Kenya, PO Box 40658, Nairobi, Kenya

† Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

The important early Miocene site of Buluk, Northern Kenya¹, was re-surveyed in 1983. As a result, the faunal list for this site has been expanded and, as we report here, now includes three hominoid and at least one cercopithecoid species. Some of the species were known previously only from North African sites, but others are typically East African. Only *Proconsul* species have been found before at other East African Miocene sites of equivalent age, but the largest hominoid species from Buluk is an early species of *Sivapithecus*.

The Buluk fossil site (36° 36' E, 4° 16' N) is a small exposure of claystones with coarse sandstone and conglomerate channel fills belonging to the Buluk Member of the Miocene Bakate Formation²⁻⁴. The fossils are found almost exclusively in the channel deposits and are often associated with silicified wood. K/Ar dates on a basalt just above the fossiliferous sediments show that the fossils are older than 17.2 Myr⁴. The 1983 expedition collected new specimens and the faunal list previously given by Harris and Watkins¹ has been expanded and consolidated (see Table 1). Of particular importance is the presence of *Libycochoerus massai* and the crocodile *Tomistoma*, known before only from North Africa. Some fossil fruits were found which

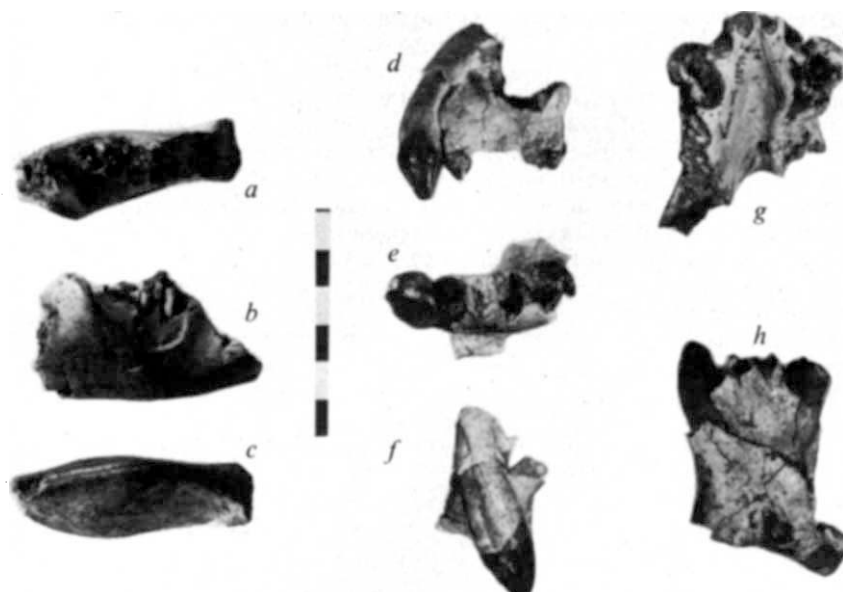


Fig. 1 a-c, Occlusal, right lateral and inferior views, respectively, of KNM-WS 124 (right mandible). d-f, Left lateral, occlusal and anterior views, respectively, of maxilla, KNM-WS 11599. g, h, Occlusal and anterior views, respectively, of mandible, KNM-WS 11599.

are similar to those reported from Rusinga Island as belonging to the form genus *Anonasperrum*⁵.

Three hominoid species are now known from Buluk (see Table 2). A small hominoid about the size of *Micropithecus clarki*⁶ is represented by KNM-WS 12610, a weathered mandible with very worn M_{2-3} . A medium-sized species is known by a right mandibular body with the roots of M_{2-3} and an isolated right M^1 (KNM-WS 12602 and 12608). These remains are about the size of the type specimen of *Kenyapithecus wickeri* from Fort Ternan⁷. On the basis of size there is no doubt that the middle-sized form is a separate species from the better represented large hominoid.

The large hominoid is known by gnathic fragments. The two most complete specimens are described here and illustrated in Fig. 1. KNM-WS 124 is a right mandibular body preserved from the midline at the symphysis to a break behind the M_3 roots.

The crowns of the molars are absent and M_1 was lost in life, as judged from the lack of roots and a large patch of alveolar resorption beneath the distal position of the tooth and beneath M_2 . The base at the symphysis is thickened so that the mental foramen lies near the edge of a hollowed surface under the premolars. The body is deepest under M_2 (38 mm) and is 19 mm thick at this point. The lingual surface is relatively flat, but hollowed in the middle behind M_3 . The buccinator groove is wide (~10 mm) and the take-off of the ramus is twisted so that its surface faces laterally and posteriorly. The external buttress continuing from the anterior edge of the ramus swings in a smooth arc under the molar roots and beneath the mental foramen to continue as the canine ridge. The anterior digastric rugosity is faint. The lengths of M_2 and M_3 have been estimated from their roots at 13 and 15 mm, respectively. This specimen is remarkably similar in parts preserved, size and morphology to GSP 11706 from the Potwar Plateau, Pakistan⁸.

Specimen KNM-WS 11599 is an associated left maxilla and anterior mandible. The mandible fragment has the alveoli for the incisors, roots of the left canine and premolars and the crown of the right canine preserved. The symphysis and the left body to a level under M_1 are present. The symphysis is deep. The alveolar plane angles posteriorly at ~45° and is highly excavated, rising steeply on either side from a midline sulcus. There are clearly defined superior and inferior tori, the latter

Table 1 Faunal list

Pisces	Clariidae <i>Protopterus</i> sp.
Reptilia	Pelomedusidae Trionychidae <i>Crocodylus</i> cf. <i>niloticus</i> <i>Tomistoma</i> sp.
Mammalia	<i>Prohylobates</i> sp. cf. <i>Micropithecus clarki</i> cf. <i>Kenyapithecus wickeri</i> <i>Sivapithecus</i> sp. Indeterminate rodents <i>Hyainalouros nyanzae</i> Very large amphicyonine <i>Gomphotherium</i> cf. <i>angustidens</i> New genus and species of gomphothere <i>Prodeinotherium hoblely</i> <i>Megalohyrax championi</i> <i>Dicerorhinus leakeyi</i> <i>Chilotheridium pattersoni</i> <i>Aceratherium acutirostratum</i> <i>Hyoboa</i> sp. <i>Diamantohyus africanus</i> <i>Libycochoerus massai</i> <i>Lopholistriodon</i> cf. <i>moruoroti</i> <i>Canthumeryx</i> cf. <i>sirtensis</i> <i>Dorcatherium chappuisi</i> <i>Dorcatherium pigotti</i>

Table 2 New higher primates from Buluk

cf. <i>Micropithecus</i>	
KNM-WS 12717	Mandible with M2-3
KNM-WS 12636	Right femur
cf. <i>Kenyapithecus wickeri</i>	
KNM-WS 12602	Right mandible
KNM-WS 12607	Upper right M3
<i>Sivapithecus</i> sp.	
KNM-WS 124	Right mandible
KNM-WS 125	Left mandible
KNM-WS 11599	Associated mandible and left maxilla with lower right C, upper C, P3, M2
KNM-WS 12601	Right mandible with I and M2 germs
KNM-WS 12606	Lower left M3
KNM-WS 12608	Proximal phalanx
Indeterminate hominoids	
KNM-WS 12603	Vertebral body
KNM-WS 12604	Head and neck of femur
KNM-WS 12605	Femoral head

being especially well-developed and curving smoothly round to the base. The canine crown is tetrahedral with a small apical dentine exposure; it is rotated so that its long axis lies obliquely. The maximum length is 18.3 mm and the width 13.0 mm. The canine roots are very long and converge inferiorly. The body is distinctly hollowed beneath the premolars and behind the curving canine ridge. The left maxilla contains the canine, P³ and M² crowns and the alveoli and roots of P⁴ and M¹. It reaches to the edge of the incisive foramen medially. This is at a level between the premolars. The upper canine is also long-rooted and stubby-crowned; its root defines a strong canine pillar. The zygomatic process swings abruptly laterally so that there is a deep canine fossa. The canine is 16.5 mm mesiodistally by 14.6 mm mediolaterally, with a height of 15 mm. The P⁴ is bicuspid with a large and high buccal cusp. The remaining buccal half of the M² is hardly worn, has 1.5-mm-thick enamel over the distolingual cusp and a small contact facet for M³. The maxillary sinus extends as a single cavity as far anteriorly as M¹ and hardly evaginates between the molar roots.

On the basis of several features, we believe that this material is distinct from the well-known African early Miocene *Proconsul* species⁹⁻¹². Those features which distinguish it from *Proconsul* are those which compare most favourably with *Sivapithecus*. The most obvious of these are the presence of two mandibular tori, the absence of cingula on the teeth, the strong outward rotation of the canines, the short, deep and strongly buttressed facial skeleton, the deep canine fossa and the posterior position of the incisive foramen.

On the basis of the parts known, it is difficult to offer a differential diagnosis that clearly separates the Buluk material from the genus *Sivapithecus*. A number of authorities have also included middle Miocene hominoids from Europe and Africa in this genus¹³⁻¹⁶. Unfortunately, there seems to be little consensus at present as to the exact status of the species in this genus, the hypotheses of most workers being different. Under these circumstances we believe it to be most useful to assign the large Buluk species to *Sivapithecus*. There is a close resemblance to larger specimens from India and Pakistan which have been placed in *Sivapithecus indicus*⁸ and from Turkey which have been placed in *Sivapithecus metei*¹³ or *S. indicus*¹⁷. *Sivapithecus* as presently being used by many workers almost certainly includes many widely dispersed species of early to middle Miocene hominoids which are similar in jaw and tooth morphology and which are representatives of a long-lived radiation, but whose exact cladistic relationships are not yet clear. Members of this genus were present in Africa over 17 Myr ago and persisted at least in Asia until about 7.5 Myr ago¹⁸. The Buluk species of *Sivapithecus* extends the time range of this genus back into the early Miocene. The similar and well-known *S. indicus* from the U level of the Potwar Plateau¹⁹ at close to 8 Myr is only half as old as the Buluk species. A single fragment of maxilla has previously been assigned to this genus⁹, but this specimen's provenance in East Africa is unknown. It is the type specimen of *Sivapithecus africanus*⁹ and it could be as old as 18 Myr if it is, as originally thought, from Rusinga Island. Others have suggested that the specimen came from Maboko²⁰, which would date it roughly equivalent to Fort Ternan, a site from which were obtained fragmentary specimens which have also been assigned to this genus¹⁶. As recently as 1978 this maxilla was placed in *Proconsul nyanzae* by the same authorities who now call it *Sivapithecus*²¹.

The unequivocal presence of *Sivapithecus* in Africa over 17 Myr ago also gives a new slant to the debate on hominoid origins. The presence of putative *Pongo* ancestors¹³ at 11–8 Myr in Asia was thought to be consistent with divergence dates between *Pongo* and African hominoids calculated from molecular evidence²². If there is indeed a unique ancestral relationship between the Buluk species and the later Asian ones, the fossil and molecular evidence are no longer congruent because the divergence time would be twice as great.

The recognition of *Pongo*-like features in Asian *Sivapithecus* species²² has led many authorities to place all similar fossil

species in the same clade and, by so doing, removed them from candidacy as either African ape or human ancestors. We believe this to have been premature. Leaving aside the possibility that the Buluk species represents an extinct lineage completely independent of Asian *Sivapithecus*, it is more likely in our opinion that the *Sivapithecus* morphology was present in the early ancestors of all great apes and hominids. It has become increasingly clear that is difficult to recognize primitive or derived features in these hominoids, one authority recently uniting *Sivapithecus*, *Pongo* and *Homo* in a single clade distinct from the African ape clade²³. Rather than conclude that the Buluk *Sivapithecus* represents the beginning of the *Pongo* clade in Africa over 17 Myr ago, we view it as an early species of a genus from which both African apes and hominids could be derived. The existence of such an old specimen of *Sivapithecus* does not necessarily say very much about the divergence time of the African and Asian hominoid clades.

We thank the Government of Kenya and the Governors of the National Museums of Kenya. The research was financed by the National Geographic Society of Washington, DC, the Garland Foundation and the National Museums of Kenya. We thank Bw. Kamoya Kimeu and his prospecting team for their invaluable help. Drs M. G. Leakey, I. McDougall, D. R. Pilbeam and R. Watkins helped in many ways.

Received 13 March; accepted 30 September 1985.

- Harris, J. M. & Watkins, R. *Nature* **252**, 576–577 (1974).
- Watkins, R. thesis, Univ. London (1982).
- Watkins, R. *Proc. geol. Soc. Lond.* (in the press).
- McDougall, I. & Watkins, R. *Nature* **318**, 175–178 (1985).
- Chester, K. I. M. *Palaeontographica* **101**, 30–71 (1957).
- Fleagle, J. *Folia Primatol.* **24**, 1–15 (1975).
- Andrews, P. & Walker, A. in *Human Origins* (eds Isaac, G. L. & McCown, E. R.) 274–304 (Benjamin, Menlo Park, 1976).
- Pilbeam, D., Rose, M. D., Badgley, C. & Lipshutz, B. *Postilla* **181**, 1–94 (1980).
- Clark, W. E., LeGros, & Leakey, L. S. B. *Fossil Mammals of Africa* **1**, 1–117 (1951).
- Pilbeam, D. *Bull. Peabody Mus. nat. Hist.* **31**, 1–185 (1969).
- Whybrow, P. J. & Andrews, P. *Folia Primatol.* **30**, 115–125 (1978).
- Andrews, P. *Bull. Br. Mus. nat. Hist.* **A30**, 85–224 (1978).
- Andrews, P. & Tsekaya, I. *Palaeontology* **23**, 85–95 (1980).
- De Bonis, L. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 625–649 (Plenum, New York, 1983).
- Greenfield, L. O. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 695–703 (Plenum, New York, 1983).
- Kay, R. F. & Simons, E. L. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 577–624 (Plenum, New York, 1983).
- Simons, E. L. *Phil. Trans. R. Soc. B292*, 21–41 (1981).
- Tauxe, L. *Nature* **282**, 399 (1979).
- Pilbeam, D. *Nature* **295**, 232–234 (1982).
- Andrews, P. J. & Molleson, T. I. *Bull. Br. Mus. nat. Hist.* **A32**, 19–23 (1979).
- Simons, E. L., Andrews, P. & Pilbeam, D. R. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 120–146 (Harvard University Press, 1978).
- Andrews, P. & Cronin, J. E. *Nature* **297**, 541–546 (1982).
- Schwartz, J. H. *Nature* **308**, 501–505 (1984).

Age of hominoid-bearing sequence at Buluk, northern Kenya

Ian McDougall* & Ronald T. Watkins†

* Research School of Earth Sciences, Australian National University, PO Box 4, Canberra, ACT 2601, Australia

† Department of Geology, University of St Andrews, College Gate, St Andrews, Fife, Scotland KY16 9ST, UK

In northern Kenya, detrital terrigenous sediments of the Buluk Member of the Bakate Formation contain an important vertebrate fauna, which includes material from three hominoid species, one of which has been assigned to the genus *Sivapithecus* and is described elsewhere in this issue¹. Here we report K–Ar age data on feldspars from volcanic rocks in the area that show the Buluk Member sediments were deposited between 16 and something more than 17.2 Myr ago in the late early Miocene.

The area from which the fossils were recovered lies in the upper Bakate Valley, 45 km east of Ileret and Lake Turkana in Kenya (Fig. 1). The geology has been described previously^{2,3}. Briefly, the sequence dips ~6° to the south-west (Fig. 2), and

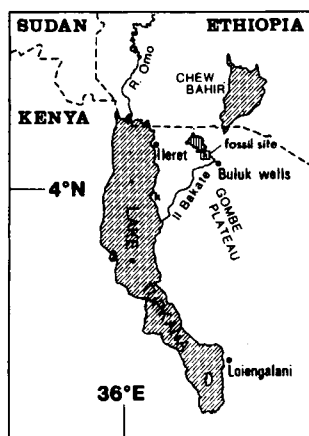


Fig. 1 Map showing location of area in northern Kenya.

consists of volcanic rocks, volcanoclastic sediments and normal detrital sediments; a composite stratigraphical section is given in Fig. 3. The Bakate Formation comprises basalts of the Irile Member, overlain by the mainly sedimentary Buluk Member, in turn overlain by basalts of the Il Jimma Member, and is succeeded disconformably by volcanoclastic sediments and ignimbrites of the Gum Dura Formation. The Buluk Member in its type section, near the hominoid locality (Fig. 2), is ~50 m thick. The lower 30 m of the Member consists predominantly of soft, poorly indurated clays, silts, sands and conglomerates of fluvial aspect. Vertebrate fossils have been recovered from sandstones, sandy siltstones and pebble conglomerates over ~15 m of the section (Fig. 3). Locally within the Buluk Member, just above the vertebrate bearing sediments, there is an essentially aphyric basalt flow ~2 m thick. The upper part of the Buluk Member, up to 20 m thick, consists of redeposited air-fall pantelleritic tuffs which are remarkably free of detrital components.

Samples for isotope dating were collected from several volcanic units in the sequence (Fig. 3) so that the reliability of the measurements could be assessed by comparison with the known stratigraphical order. Unfortunately, no suitable samples of basalt from the Irile Member were available for dating.

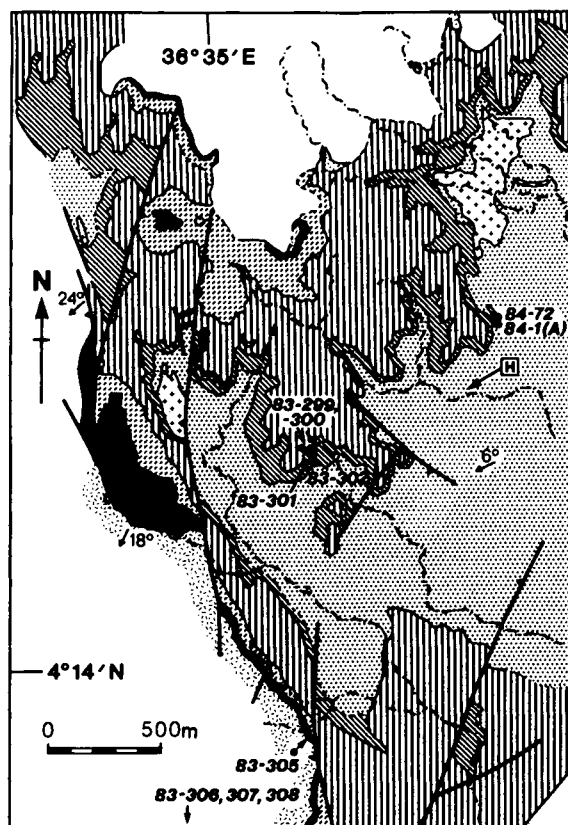


Fig. 2 Simplified geological map of the Buluk vertebrate fossil locality showing collection sites of samples used for dating (●) and of the hominoid specimens (H). Stratigraphical division and key is given in Fig. 3.

In all cases, separation of high-temperature volcanic feldspar provided suitable material for isotope dating by the conventional K-Ar method. Alkali feldspar of high purity was obtained from ignimbrite samples of the Gum Dura Formation and two tuff samples from the upper part of the Buluk Member, using

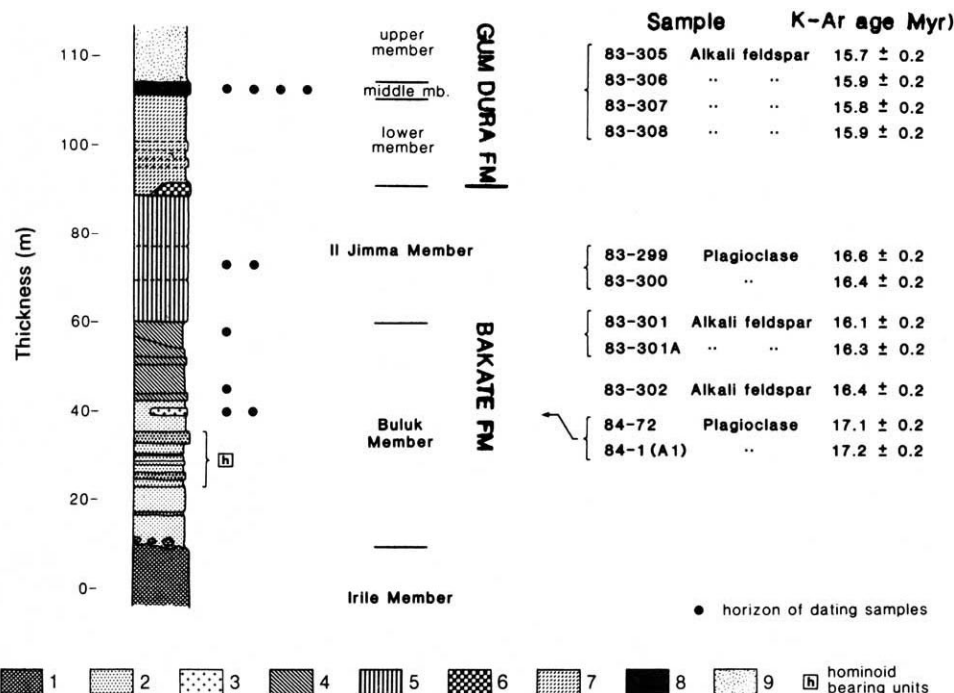


Fig. 3 Generalized stratigraphic section showing the units exposed at the Buluk fossil locality, the position of the dated samples in the sequence and their measured ages. 1, Deeply weathered porphyritic feldspar-phyric basalt lavas; 2, fluvial claystones, siltstones, sandstones and conglomerates of terrigenous origin; 3, thin aphyric or finely porphyritic basalt lava flows; 4, water redeposited air-fall pantelleritic tuffs; 5, fine-grained aphyric basalt lavas; 6, localized hydrothermal deposits of travertine and sinter; 7, pyroclastic pantellerites including air-fall deposits (largely reworked), poorly welded ignimbrites, and volcanic breccias; 8, strongly welded pantelleritic ignimbrite; 9, generally poorly welded pantelleritic ignimbrites.

Table 1 K-Ar ages on feldspars from ignimbrites, tufts and basalts from the Buluk area, northern Kenya

Sample no.	Sample	Grid reference HBH-	K (wt %)	Radiogenic ^{40}Ar (10^{-10} mol g $^{-1}$)	100 rad. ^{40}Ar	Calculated age (Myr $\pm$ 1 s.d.)	Average age (Myr $\pm$ 1 s.d.)
					Total ^{40}Ar		
Ignimbrites of Gum Dura Formation							
83-305	Alkali feldspar	985677	5.795, 5.773	1.571	85.9	15.6 $\pm$ 0.2	15.7 $\pm$ 0.2
				1.594*	84.7	15.8 $\pm$ 0.2	
				1.582*	80.4	15.7 $\pm$ 0.2	
83-306	Alkali feldspar	958668	5.892, 5.910	1.627	95.5	15.8 $\pm$ 0.2	15.9 $\pm$ 0.2
				1.635*	81.2	15.9 $\pm$ 0.2	
83-307	Alkali feldspar	979656	5.862, 5.880	1.605	98.6	15.7 $\pm$ 0.2	15.8 $\pm$ 0.2
				1.629*	83.5	15.9 $\pm$ 0.2	
83-308	Alkali feldspar	982640	5.826, 5.800	1.599	97.0	15.8 $\pm$ 0.2	15.9 $\pm$ 0.2
				1.613*	80.2	15.9 $\pm$ 0.2	
Basalt, Il Jimma Member, Bakate Formation							
83-299	Plagioclase	986690	1.676, 1.686	0.487	41.2	16.6 $\pm$ 0.2	16.6 $\pm$ 0.2
				0.486	43.5	16.6 $\pm$ 0.2	
				0.488*	39.9	16.6 $\pm$ 0.2	
83-300	Plagioclase	985690	1.690, 1.685	0.483	48.3	16.4 $\pm$ 0.2	16.4 $\pm$ 0.2
				0.480*	41.5	16.3 $\pm$ 0.2	
Tufts of Buluk Member, Bakate Formation							
83-301	Alkali feldspar	987689	6.512, 6.506	1.834	95.6	16.2 $\pm$ 0.2	16.1 $\pm$ 0.2
				1.819	92.0	16.1 $\pm$ 0.2	
				1.827*	85.5	16.1 $\pm$ 0.2	
83-301A	Alkali feldspar	987689	6.435, 6.403	1.822	84.2	16.3 $\pm$ 0.2	16.3 $\pm$ 0.2
				1.821	88.9	16.3 $\pm$ 0.2	
83-302	Alkali feldspar	987689	5.734, 5.725	1.639	90.8	16.4 $\pm$ 0.2	16.4 $\pm$ 0.2
				1.626	84.0	16.3 $\pm$ 0.2	
Basalt flow, Buluk Member, Bakate Formation							
84-72	Plagioclase	994696	1.133, 1.141	0.342	83.0	17.2 $\pm$ 0.2	17.1 $\pm$ 0.2
				0.335*	77.8	16.9 $\pm$ 0.2	
84-1 (Al)	Plagioclase	994696	1.303, 1.307	0.393	78.1	17.3 $\pm$ 0.2	17.2 $\pm$ 0.2
				0.391	80.4	17.2 $\pm$ 0.2	

$\lambda_e + \lambda_e' = 0.581 \times 10^{-10} \text{ yr}^{-1}$, $\lambda_\beta = 4.962 \times 10^{-10} \text{ yr}^{-1}$, $^{40}\text{K}/\text{K} = 1.167 \times 10^{-4} \text{ mol mol}^{-1}$.

Grid reference to East Africa Grid. Buluk 1:100,000 map, series Y633, edition 1, 1960, Survey of Kenya. Mass of sample used in Ar extraction normally 0.30 g for alkali feldspar and 0.50 to 0.70 g for plagioclase.

* Ar extraction done together with 1.0 g of essentially zero age basalt as a flux.

standard heavy liquid and magnetic separation techniques. No detrital contamination was seen in the tufts. All the alkali feldspars were fresh and limpid, with no evidence of unmixing.

Basalt samples collected from the flow within the Buluk Member and from the Il Jimma Member were essentially aphyric, although the Buluk basalt contained sparse plagioclase and altered olivine microphenocrysts. Unaltered plagioclase as thin laths, averaging 0.1–0.2 mm in length, is the most important constituent of the basalts, together with lesser amounts of pyroxene, iron oxide, some olivine and up to ~10% of green mineraloid, clay or chlorite. Because of the presence of alteration products, the samples were judged to be unsuitable for dating as whole rocks, and instead the plagioclase was separated, albeit with difficulty because of its small grain size. To obtain sufficiently good separates, the concentrates were treated briefly in dilute HF to remove mineraloid and clay in composites with the feldspar. The grain size of the plagioclase separates averaged ~50 μm .

The K-Ar age measurements were made according to procedures described previously^{4,5} using flame photometry for K and isotope dilution for Ar. The Ar isotope measurements were carried out in a VG-Isotopes MM1200 mass spectrometer operated in the static mode online with a computer. As noted earlier⁵, it is difficult to extract Ar quantitatively from volcanic alkali feldspars. Each sample, contained within a Mo crucible in a water-cooled silica glass bottle attached to the high-vacuum extraction line, was held at ~1,600 °C for 40–60 minutes using radiofrequency heating. In some cases it was suspected that there was incomplete Ar recovery, despite the high temperature

of extraction. Consequently, some replicate runs were made using 1.0 g of a very young basalt as a fluxing agent with the feldspar, following the suggestion of McDowell⁶. The basalt (sample no. 83-229) used as flux contained $<7.0 \times 10^{-14} \text{ mol g}^{-1}$ radiogenic Ar. Quantitative extraction of Ar from plagioclase seems to be much less of a problem compared with alkali feldspar.

Results are given in Table 1, arranged in stratigraphical order. Reproducibility of both K and Ar measurements is generally better than 1%, so that the overall precision is ~1%. The average age for each sample is reported with a minimum error of 1%, as the combined uncertainties in the calibration of the Ar tracer and K standards are at this level. Several of the earlier Ar results for alkali feldspars have been omitted from Table 1, as replicate measurements clearly showed that there had been incomplete extraction of Ar. In one case where this was suspected, the crucible was heated again at an even higher temperature ($>1,700$ °C) for a further 45 min, and additional radiogenic Ar was detected, ~4% of the total radiogenic Ar. Although the four alkali feldspars separates from the Gum Dura ignimbrites averaged 1.0% higher radiogenic ^{40}Ar when the basalt flux was used compared with the results without the flux, we regard the agreement as satisfactory.

The consistency of results from samples from the same or closely related stratigraphical levels (Table 1), together with an overall decrease in age upwards through the sequence from 17.2 ± 0.2 to 15.8 ± 0.2 Myr (Fig. 3), provides confidence that the measured ages closely reflect the time since crystallization of these volcanic rocks and the deposition of the associated detrital

sediments. However, detailed examination of the results reveals some minor inconsistencies that must be addressed.

Alkali feldspar separated from four samples of ignimbrite from the middle member of the Gum Dura Formation yield completely concordant ages, with a mean of 15.8 ± 0.2 Myr (Table 1). The four samples were collected at different localities extending as far as 4 km south of the map area; only the most northerly locality is shown in Fig. 2. The member comprises two welded ignimbrite sheets, and the fact that the samples are from the same or closely related welded ignimbrite units is supported by the concordant ages and the similarity of the K contents of the alkali feldspars. Because of the unaltered nature of the feldspars, the excellent Ar retention properties of volcanic alkali feldspar and the concordancy of the results, we are confident that the measured age is that of crystallization and cooling of the ignimbrites. Thus, 15.8 ± 0.2 Myr is considered a firm younger limit for the age of the Bakate Formation. Ages determined in this study are younger than those of 16.5 ± 0.1 and 16.6 ± 0.1 Myr (recalculated using decay constants used here), derived from $^{40}\text{Ar}/^{39}\text{Ar}$ step-heating experiments by Fitch *et al.*⁷ for alkali feldspar from the two ignimbrites from the middle member of the Gum Dura Formation elsewhere in the area.

Plagioclase from two separate samples of basalt of the II Jimma Member, collected 50 m apart laterally and probably from the same massive flow, yielded concordant ages of 16.6 ± 0.2 and 16.4 ± 0.2 Myr (Table 1, Fig. 3). Taken at face value, these results suggest the basalt erupted 16.5 ± 0.2 Myr ago, consistent with the age determined for the overlying Gum Dura Formation. Fitch and Miller⁸ reported a K-Ar age of 17.7 ± 1.4 Myr (recalculated) on a whole rock sample of basalt from the II Jimma Member; the large uncertainty assigned means that the age is indistinguishable from the more precise age given here.

Two separate alkali feldspar concentrates (83-301 and 83-301 A) from a coarse tuff of the Buluk Member, collected ~2 m below the II Jimma basalt, gave ages of 16.1 ± 0.2 and 16.3 ± 0.2 Myr, respectively. Another alkali feldspar separated from a tuff ~8 m lower stratigraphically and ~2 m above the red claystones of the Buluk Member, yielded a K-Ar age of 16.4 ± 0.2 Myr (Table 1). Although the average of these ages is slightly younger than the mean age measured on the overlying II Jimma basalts, the difference between them is small and not significant when the errors are taken into account. We conclude from these virtually concordant results that the tuffs of the upper part of the Buluk Member and the basalts of the overlying II Jimma Member are indistinguishable in age, estimated as 16.4 ± 0.2 Myr.

Plagioclase concentrates from two separate samples of basalt from the same lava flow ~2 m thick within the terrigenous sediments of the Buluk Member, yielded concordant K-Ar ages of 17.1 ± 0.2 and 17.2 ± 0.2 Myr (Table 1, Figs 2, 3). Note that the K contents of the two concentrates differ appreciably, but that the ages agree. We believe the measured age is recording the time of crystallization and cooling of the basalt. Acceptance of this age suggests that >0.5 Myr elapsed between eruption of the basalt flow and the silicic volcanic eruptions that produced the tuffs of the upper part of the Buluk Member. The vertebrate-bearing sediments of the Buluk Member lie just below the basalt (Fig. 3), and as no significant breaks are recognized in the sequence we believe the underlying sediments are unlikely to be very much older.

Thus, the 70-m of sequence sampled yielded K-Ar ages on volcanic feldspars that are generally consistent with the stratigraphy, and in the range 17.2–15.8 Myr, covering a time span of <1.5 Myr (Fig. 3) in the late early Miocene⁹. Vertebrate bearing sediments of the Buluk Member are only a little older than the basalt dated at 17.2 ± 0.2 Myr, and probably no older than 18 Myr. The ages reported here also provide useful data for establishment of a time framework for the stratigraphical sequence found in this part of Kenya².

We thank the National Museums of Kenya and the Government of Kenya for facilitating our field studies, and R. E. F. Leakey for support in the field. R.T.W. was supported by

research grants from National Environmental Research Council. Technical support for the K-Ar dating was provided by M. Cowan, T. Davies, R. Maier and R. Rudowski.

Received 2 April; accepted 17 July 1985.

1. Leakey, R. E. F. & Walker, A. *Nature* **318**, 173–175 (1985).
2. Watkins, R. T. thesis Univ. London (1983).
3. Harris, J. M. & Watkins, R. T. *Nature* **252**, 576–577 (1974).
4. McDougall, I. & Schmincke, H.-U. *Bull. Volcan.* **40**, 57–77 (1977).
5. McDougall, I., Maier, R., Sutherland-Hawkes, P. & Gleadow, A. J. W. *Nature* **284**, 230–234 (1980).
6. McDowell, F. W. *Isotope Geosci.* **1**, 119–126 (1983).
7. Fitch, F. J., Watkins, R. T. & Miller, J. A. *Nature* **254**, 581–583 (1975).
8. Fitch, F. J. & Miller, J. A. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y., Howell, F. C., Isaac, G. Li. & Leakey, R. E. F.) 123–147 (University of Chicago Press, 1976).
9. Harland, W. B. *et al.* *A Geologic Time Scale* (Cambridge University Press, 1982).

Temporal hyperacuity in the electric sense of fish

Gary Rose & Walter Heiligenberg

Neurobiology Unit, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, USA

It has recently become evident that sensory thresholds for certain tasks are lower than those expected from the properties of individual receptors¹. This perceptual capacity, termed hyperacuity, reveals the impressive information-processing abilities of the central nervous system. Although much is known about spatial hyperacuity, temporal hyperacuity has received little attention². Here we demonstrate that an electric fish, *Eigenmannia*, can detect modulations in the timing (phase) of an electrical signal at least as small as 400 ns. Such sensitivity exceeds the temporal resolution of individual phase-coding afferents. This hyperacuity results from a nonlinear convergence of parallel afferent inputs to the central nervous system; subthreshold inputs from particular areas of the body surface accumulate to permit the detection of these extremely small temporal modulations.

Behavioural experiments have demonstrated that certain species of electric fish can perform remarkable analyses of the temporal structure of electrical signals^{3–5}. These animals produce an electrical signal within a species-specific frequency range via an electric organ, and they detect these signals by electroreceptors located throughout the body surface. In the context of one electrosensory behaviour, the jamming avoidance response (JAR), the fish *Eigenmannia* determines whether a neighbour's electric organ discharge (EOD), which is jamming its own signal, is higher or lower in frequency than its own. The fish then decreases or increases its frequency, respectively. To determine the sign of the frequency difference, the fish must detect the modulations in the amplitude and in the differential timing, or temporal disparity, of signals received by different regions of its body surface^{6,7}. The fish is able to shift its discharge frequency in the appropriate direction in at least 90% of all trials for temporal disparities as small as 400 ns (Fig. 1).

Intracellular electrophysiological measurements show that the phase-locked responses of even the best afferent recorded are too jittery to permit such fine temporal resolution⁸. The jitter of this time-locking can be described by its standard deviation. Even the most accurate phase-coders time-lock their spikes with a standard deviation of ~10 μ s (Fig. 2). For a sample period of 300 ms (and thus ~100 EOD cycles), which is the latency of the JAR, the 95% confidence intervals around the mean phase of occurrence of such an afferent's spikes are 2.0 μ s. Yet the fish is able to detect time disparities of several hundred nanoseconds⁸. Statistically, it would appear to be impossible for the fish, using only the information gathered from any single

Fig. 1 Percentage of correct (that is, positive) jamming avoidance responses (JARs) compared with maximal available temporal disparity. Data from three fish are presented. Inset, averaged JARs ($n=10$) for S_2/S_1 ratios of 0.1 (top) and 0.001 (bottom), corresponding to maximal temporal disparities of 80 μs and 800 ns, respectively. Arrows indicate points at which the sign of the 4-Hz frequency difference was changed. **Methods.** An immobilized fish was placed in a two-compartment chamber, and respiration by a constant flow of water through the mouth. The fish's body rested in a slot, lined with silicone grease, thus providing an electrical isolation of >40 dB between compartments (that is, a signal delivered to one compartment would be present in an adjacent compartment at a level less than $1/100$ th). The fish's own electric organ discharge was eliminated by intramuscular injection of 1–2% Flaxedil and replaced by a sinusoidal signal (S_1) which matched the fish's natural signal in frequency and amplitude. This EOD-replacement signal and the jamming signal (S_2) were presented through the same electrodes located in the head compartment; positive pole in the fish's mouth, negative poles in the water on either side of the fish. S_1 alone was presented in the posterior (tail) compartment; positive pole in the dorsal musculature, negative poles in the water. Thus, phase modulations existed between the signals in these two compartments. With f_1 and f_2 being the frequencies of S_1 and S_2 , a rise in the amplitude of the signal in the head compartment was accompanied by a phase advance or delay of this signal, relative to the signal in the tail compartment, when $f_2 < f_1$ or $f_2 > f_1$, respectively. Since S_1 and S_2 were delivered through the same electrodes, their amplitude ratio (S_2/S_1) was constant at all points in the chamber and equal to the ratio of the outputs of the two function generators that produced them. Thus, the degree of amplitude and phase modulations experienced by the electroreceptors in the head region could be precisely determined. The maximal time disparity, DT , is calculated from the following formula: $DT = aT/2\pi$, where T is the length of the period of S_1 , the signal used to replace the fish's electric organ discharge; S_2 is the jamming signal and a is the amplitude ratio between S_2 and S_1 . While this stimulation arrangement had the advantage of permitting exact determination of the available temporal disparity, it is a less effective means of eliciting the JAR than the multi-chamber arrangement (Fig. 3) or free field conditions (no compartments). The threshold value determined should therefore be considered a conservative estimate. JARs were measured according to the following procedure^{6–8}. The electric organ pacemaker was determined by recording the spinal volley signal from the electric organ pacemaker via a suction electrode at the tail. This signal was amplified 1,000 times and each discharge triggered a standard square pulse that was fed into a 100-kHz computer-controlled clock for timing analysis. The fish's pacemaker (or would-be electric organ discharge) frequency was then computed. The sign of the 4-Hz frequency difference between the jamming signal (S_2) and the replacement signal (S_1) was changed every 25.6 s by the computer. Positive JAR values indicate a correct JAR (that is, the fish shifted its discharge frequency (DF) in the direction that was appropriate for the stimulus condition at the time); upward shifts for a jamming frequency that was lower and downward shifts for a jamming frequency that was higher. Shifts in frequency were integrated over the period of stimulation, and the mean shift was taken as the JAR measurement.

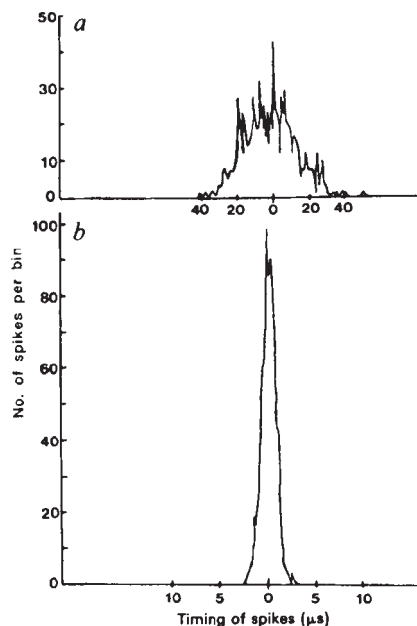
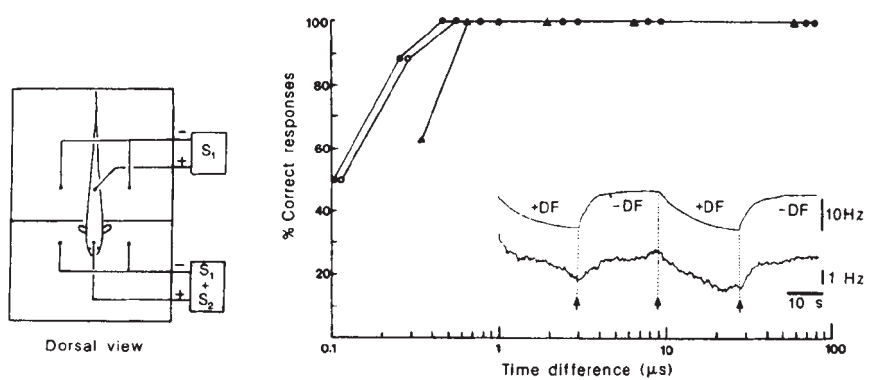


Fig. 2 *a*, Distribution of spikes around a mean time of occurrence (zero) for a single primary afferent, the most accurate phase-coder recorded. 1,024 spikes were sampled. Bin width is 0.1 μs . *b*, Electronically generated 'spike-mimics' were recorded by the same equipment used for the neural recordings. The standard deviation of the time of occurrence of these spike-mimics, with reference to the time of their production by the waveform generator, was ~ 700 ns. The neural jitter is, therefore, not a consequence of limitations in the accuracy of measuring spike occurrences.

Methods. Spikes and spike-mimics were recorded with 3 M KCl-filled micropipettes (20 M Ω). Neural recordings were made from the anterior lateral line ganglion of anaesthetized fish. For these recordings, a sinusoidal signal, matching the fish's EOD in frequency and amplitude, was delivered through a positive electrode in the fish's mouth and a negative electrode at the tail. Forty afferents were recorded. The spike-mimic was presented through electrodes in the water. When recorded by the micropipette, these spikes had a similar rise time to that of the steepest portion of the neural spikes ($1 \text{ mV } \mu\text{s}^{-1}$). The membrane potential of the neurone was -80 mV , spikes were 60 mV. Baseline voltage fluctuations were $\sim 0.5 \text{ mV}$. This large signal-to-noise ratio reduced the temporal jitter resulting from baseline-voltage fluctuations to well below that commonly experienced in extracellular recordings. The time of occurrence of spikes with reference to each synch pulse from the signal generator, defined as the moment when the steepest portion of the signal passed a trigger level, was measured via a computer-controlled clock with a resolution of 0.1 μs .

afferent, to reliably shift its frequency in the correct direction when the maximal temporal disparity available is only several hundred nanoseconds.

These findings lead to the prediction that the behavioural threshold should be higher when only a small group of receptors is stimulated, and that hyperacuity results from the convergence, within the central nervous system, of parallel phase-coding channels from sufficiently large areas of the body surface. These notions were tested using a paradigm which enabled us to stimulate selectively particular regions of the body surface (Fig. 3). Jamming avoidance responses could be elicited from stimulation of the body surface within any of the individual compartments, provided the ratio between the amplitude of the jamming stimulus (S_2) and the amplitude of the EOD-replacement signal (S_1) was sufficiently large. For example, an S_2/S_1

ratio of 0.05 produced a strong JAR when the stimulus was delivered exclusively to compartment 3 (Fig. 3*a*). With this ratio the fish experiences a stimulus that has a maximum relative amplitude modulation of 5% and gives rise to a maximal temporal disparity between the signal received by the left and right sides of the body of 28.4 μs . When the S_2/S_1 ratio was reduced to 0.0016, corresponding to a maximal temporal disparity of 1.8 μs , no JAR was observed on individual stimulation of compartments 2, 3 or 4 (Figs 3*b*, 4). Simultaneous stimulation of the body surface in these three compartments, however, elicited a strong JAR (Figs 3*c*, 4).

By electronically modulating amplitude and phase independently^{7,9}, one can demonstrate that the resolution of amplitude modulations alone is not the limiting factor in this experiment: the animal fails to increase its JAR to stimulus regimens

Fig. 3 Averaged JARs. A signal (S_1) replacing the fish's own electric organ discharge, and a jamming signal (S_2), which was alternately 4 Hz higher or lower than the replacement signal, were applied to individual compartments (a, b) or simultaneously to all three compartments (c). The amplitude ratio between these two signals, measured lateral to the fish's body, is given on the left-hand side of the figure. Centre (darkest) trace is the mean, traces to either side indicate the standard error.

Methods. An immobilized fish was placed in a multi-compartment chamber. Stimuli were delivered through wires in each of the five compartments. The EOD-replacement signal (S_1) was delivered to each compartment through an electrode in the dorsal musculature and an electrode at the bottom of the compartment, current at any given point in time thus being of the same polarity for all parts of the body surface. The 'jamming signal' (S_2) was delivered, in each chamber, through two electrodes straddling the fish. With this geometry of presentation, current flowed from outside to inside on one side of the body surface, and from inside to outside on the other side, the ventral surface being only minimally affected by the jamming stimulus. Maximal temporal disparity between the signal received by the ventral surface and that received by either lateral surface is DT (as computed according to the formula in Fig. 1), while the maximal temporal disparity between the signals received by the right and left sides is 2DT. The two stimuli to each compartment could be individually delivered or removed by means of a switching network located outside the chamber. By stimulating two adjacent sections of the body, the fish can obtain differential phase information not only by comparing the two sides of the body within the same chamber but also by comparing contralateral sides located in different chambers. No phase differences are available between ipsilateral sides in different compartments. If one assumes linear addition of inputs from separate compartments and a linear relation between the effect of a phase modulation and the size of the contralateral reference area, the simultaneous stimulation of three chambers should yield a total effect of three times the sum of the contributions from individual chambers. This is still far less than was experimentally observed (see Fig. 4). The results were qualitatively similar for all three fish tested.

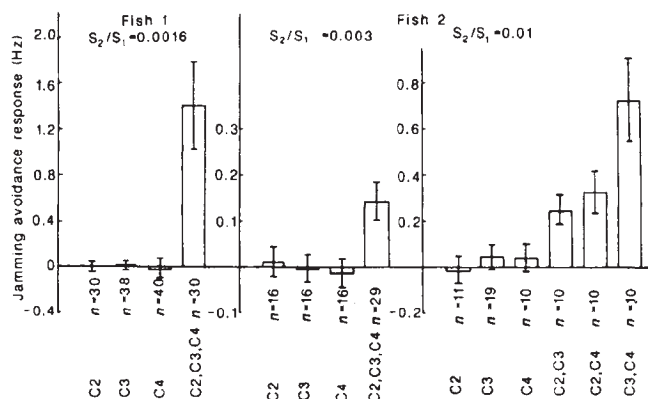
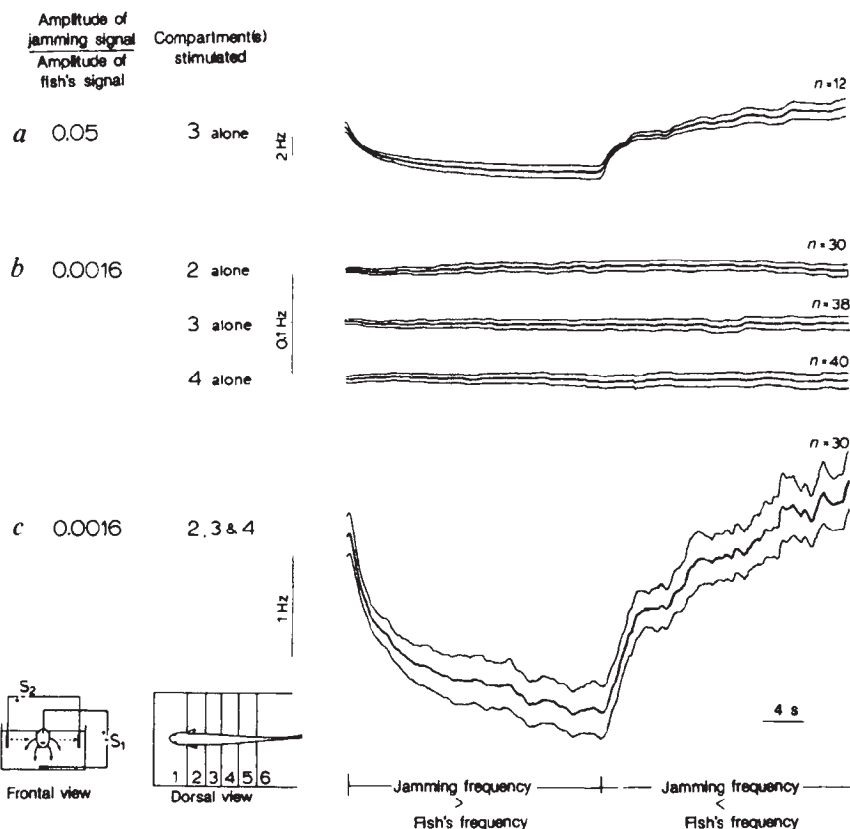


Fig. 4 Mean JAR and 99% confidence intervals for presentation of a signal (S_1) that served as a replacement for the fish's own electric organ discharge, and a jamming signal (S_2), to individual compartments (C2, C3, C4) or simultaneously to various combinations of these three compartments. The ratio between the amplitudes of these two signals is given in each case. Positive JAR values indicate that the fish shifted its electric organ discharge frequency in the appropriate direction (that is, upward when the S_2 frequency is lower than the S_1 frequency and downward for the opposite relation).

involving phase modulations at threshold level even if larger amplitude modulations are presented.

Our results indicate that electroreceptive afferent inputs converge in a nonlinear manner to achieve sensitivity to temporal disparities of several hundred nanoseconds, thus exceeding the interaural time-disparity sensitivity of owls (10 μ s; ref. 10) and of humans (15–20 μ s; ref. 11). In this process, inputs which are individually insufficient to elicit a measurable JAR (even after averaging over 2 h) produce a strong response when combined. This effect is not merely a result of the increase in total stimulus energy being presented to the fish because the amplitudes of S_1 and S_2 can be increased, while keeping their ratio constant, without altering the single-compartment response.

This temporal hyperacuity resembles hyperacuity in spatial vision in that the exquisite sensitivity is dependent on the presence of a reference point. In vision, the detection of small spatial modulations of a dot, line or edge is enhanced when a reference feature, displaced from the modulated feature, is made avail-

able¹². Similarly, for detection of small modulations in the timing of a signal, *Eigenmannia* requires a reference signal; the fish compares the timing of the signal in one region of the body surface with that in another region⁷.

This study was supported by NSF grant BNS82-05454 to W.H. and by NIH postdoctoral fellowship NS 07261 to G.R.

Received 11 July; accepted 17 September 1985.

- Westheimer, G. *Invest. Ophthalmol. vis. Sci.* **18**, 893–912 (1979).
- Simmons, J. *Science* **204**, 1336–1338 (1979).
- Bullock, T. H., Hamstra, R. H. & Scheich, H. *J. comp. Physiol.* **77**, 1–48 (1972).
- Heiligenberg, W. & Altes, R. *Science* **199**, 1001–1004 (1979).
- Hopkins, C. *Science* **212**, 85–87 (1981).
- Heiligenberg, W., Baker, C. & Matsubara, J. *J. comp. Physiol.* **127**, 267–286 (1978).
- Heiligenberg, W. & Bastian, J. *J. comp. Physiol.* **136**, 113–133 (1980).
- Carr, C. E., Heiligenberg, W. & Rose, G. *J. Neurosci.* (in press).
- Heiligenberg, W. & Rose, G. *J. Neurosci.* **5**, 515–531 (1985).
- Konishi, M., Sullivan, E. & Takahashi, T. *J. acoust. Soc. Am.* **78**(1), 360–364 (1985).
- Zwislocki, J. & Feldman, R. S. *J. acoust. Soc. Am.* **28**, 860–864 (1956).
- Westheimer, G. *Prog. Sensory Physiol.* **1**, 1–30 (1981).

Mammalian neural crest cells participate in normal embryonic development on microinjection into post-implantation mouse embryos

Rudolf Jaenisch

Whitehead Institute for Biomedical Research and
Department of Biology, Massachusetts Institute of Technology,
Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

The production of chimaeric mice by aggregating pre-implantation mouse embryos¹ or by injection of cells into the blastocyst² has been of great value in analysing the regulation of early mammalian development and in dissecting the relationships of early cell lineages^{3,4}. While the totipotent cells of the pre-implantation embryo can be grown *in vitro* and thus are readily accessible to experimental manipulation, this is not possible after the embryo has implanted into the uterus. This problem has severely hampered the analysis of cell migration and of cell lineage relationships in later stages of mammalian development. In contrast, the chicken embryo can be manipulated experimentally throughout embryogenesis and this has made the bird a favourable system for studying patterns of cell migration in the development of higher vertebrates⁵⁻⁷. In mammals, the introduction of retroviruses^{8,9} and haematopoietic cells¹⁰ has provided two means of probing post-implantation development by direct intervention. I report here that cultured neural crest cells, when microinjected into 9-day-old mouse embryos, can migrate over considerable distances and participate in normal development, and the resulting chimaeric animals show pigmentation derived from the donor cells in hair and iris. The introduction of cells into post-implantation embryos may provide the means of studying patterns of cell migration in mammalian development at a level of sophistication which so far has been restricted to the chicken system.

Neural crest cells originate from the neural folds and begin migrating into many regions of the embryo at the time of neural tube closure^{5,11}. Experiments with amphibia^{12,13} and chickens^{5,11,14,15} have shown that this embryonic population of cells gives rise to an astounding variety of different cell types, including sensory and autonomic ganglia, facial bone and cartilaginous structures, stroma of various secretory glands, and the pigment cells of the skin. When neural tubes from mid-gestation mouse embryos are explanted *in vitro*, cells which resemble morphologically the well-characterized chicken neural crest cells begin to migrate from the explant¹⁶. After several days in culture, cells begin to synthesize catecholamines and become dopa-positive, indicating their differentiation into adrenergic and prepigment cells¹⁶. Pigment formation, however, has never been observed in explants of murine, in contrast to chicken neural crest cells^{17,18}. In the present study I investigated whether explanted neural crest cells would become functional melanocytes when microinjected into developing post-implantation mouse embryos.

C57BL/6J mouse embryos were isolated at day 8.5 of gestation (5-15-somite stage) and were explanted as described in Fig. 1 legend. Cells began to migrate from the explants within 24 h and the neural tubes were detached from the dishes with tungsten needles after 2 or 3 days. Figure 1 shows a typical culture at 3 days after explantation. The migrating cells have a similar morphology to that described for neural crest cells isolated from chickens¹⁹ and mice¹⁶. Cells of neural crest morphology were seen predominantly at the edges of the migrating cells, whereas in the centres of the outgrowths extensive differentiation occurred with increasing time of culture, resulting in cells of epithelial and neuronal morphology. Further characterization of the different cell types was not attempted.

For microinjection, individual cultures were trypsinized at 3-13 days after explantation, and the cells were concentrated

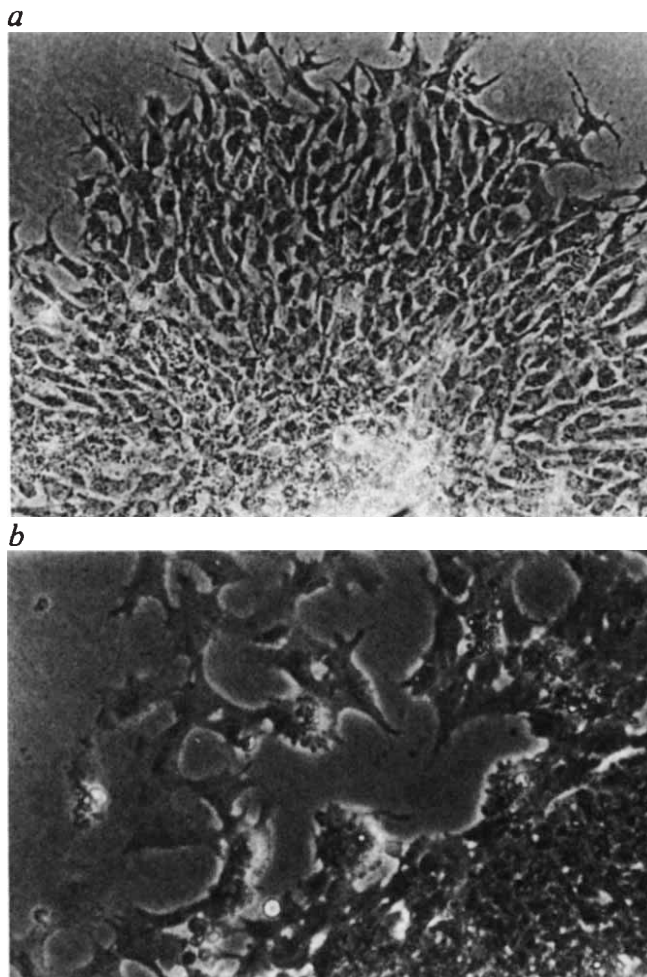


Fig. 1 *In vitro* culture of explanted neural tubes. The figure shows phase-contrast photomicrographs of cells migrating from the neural tube at 3 days after explantation. *a*, $\times 34$; *b*, $\times 112$.

Methods. C57BL/6J mouse embryos were isolated at day 8.5 of gestation (day of vaginal plug = day 0 of gestation). The embryos were freed of placenta and fetal membranes and incubated in HEPES buffer containing 1% trypsin at 5°C for 20-30 min¹⁷. Enzyme treatment was stopped by washing the embryos twice in HEPES buffer containing 10% fetal calf serum (FCS), and the neural tubes were dissected from the embryos under a dissecting microscope. The neural tubes were carefully freed of somites and all surrounding tissues and between five and eight tubes were placed in a 2-cm tissue culture dish which had been coated previously with collagen (Vitrogen) or gelatin (0.5% solution for 1 h); they were then cultured in DMEM containing 5% FCS and 5% horse serum.

by centrifugation and resuspended in 25-40 μ l of DMEM (Dulbecco's minimal essential medium) containing 10% calf serum. The total number of cells obtained per culture varied between 2×10^4 and 2×10^5 . The cells were microinjected into recipient BALB/c embryos at day 8.25, between day 8.75 and 9.25, or at day 9.75 of gestation essentially as described previously^{8,9}. Briefly, the pregnant mice were anaesthetized and laparotomy was performed by a long ventral incision. The uterus was held with a sharp forceps and cells were microinjected into individual embryos by introducing the micropipette into the ventral third of the decidual swelling. Approximately 0.1-0.5 μ l of liquid, containing 50-300 cells, was injected per embryo. Table 1 shows that approximately 60% of the injected embryos survived to adulthood.

On visual inspection of the resulting animals, I found that approximately 15% of the mice derived from embryos injected with neural crest cells between days 8.75 and 9.25 of gestation showed evidence of a pigment contribution from the donor cells

(Table 1). In contrast, none of 121 animals derived from embryos injected at day 8.25, and none of 71 animals injected at day 9.75 of gestation showed visible pigmentation. The fraction of chimaeric animals obtained, however, was strongly dependent on the age of the neural crest cell cultures. Table 2 shows that cells from 3–5-day-old cultures resulted in more than 30% chimaeric mice, while this fraction decreased to approximately 1% when cells from 11–13-day-old cultures were used for injection. This result suggests that the fraction of undifferentiated neural crest-like cells shown in Fig. 1, rather than the more differentiated cell types which accumulate with prolonged culture, give rise to pigmentation in the chimaeric mice.

Figure 2 shows typical examples of chimaeric mice. Pigmented hair was found in either the head or posterior trunk region, but never in the upper mid-trunk region or the forelimbs (see Table 3). Most chimaeras had pigmented hair in the head region which typically covered a symmetrical area extending laterally from the mid-dorsal line to both sides (Fig. 2a–c). Approximately 30% of the chimaeras were pigmented in the posterior trunk region, including the hindlimbs (7 animals; Fig. 2c, d) or the tail (3 animals; Fig. 2b). The pigmented area in the posterior trunk region frequently showed a sharp demarcation at the mid-dorsal (Fig. 2c) or mid-ventral line. However, pigmentation was not restricted to the epidermis. Microscopic inspection revealed extensive pigmentation of the iris in 10 of the 22 animals exhibiting head chimaerism. The pigment appeared to radiate in sections from the pupil to the back of the eye (Fig. 2e). Similarly, pigmented cells appeared to radiate from the optic nerve to the equator of the eye when viewed from the back (Fig. 2e). Histological examination revealed that the pigmented cells were localized in the choroid and not in the retina (U. Dräger, personal communication). This result is consistent with earlier observations indicating that the pigment of the mesodermal iris and of the choroid is derived from neural crest cells^{20,21} and not from the central nervous system (CNS), which gives rise to the pigmentation of the retina. Table 3 summarizes the types of pigment distribution seen in the 31 chimaeras. Three classes of animals were observed (types A, B and C). Pigment was restricted to the head area in 19 animals and to the posterior trunk in 9 animals. Three animals, however, exhibited pigmentation of the head as well as of the posterior trunk (type C in Table 3; compare with Fig. 2c).

The phenotype of the chimaeras indicates that neural crest cells which have been introduced into an embryo at late stages of development can participate in normal morphogenesis. How do the injected cells reach their position in the skin and iris? Because the embryo cannot be visualized at the time of injection, the pathway of donor cell migration can only be deduced from the phenotype of the chimaeras. Principally, after successful injections resulting in chimaeric mice, the cells may have been deposited either inside the embryo proper or outside it, in the extraembryonic cavities—that is, the yolk or amniotic cavity. After injection, donor cells may have migrated along aberrant pathways, for example in the trough of the exocoelomic cavity, or they may have followed normal pathways of neural crest cell migration in the dermis.

The following considerations argue against the donor cells

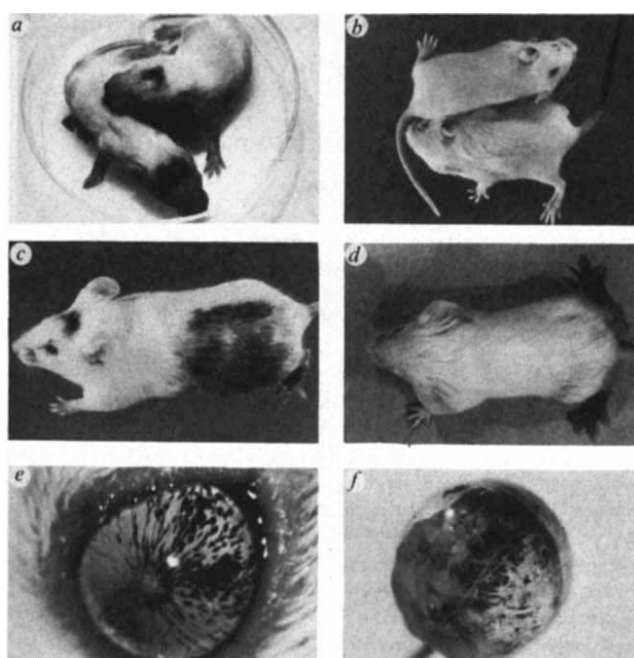


Fig. 2 Chimaeras resulting from successful injection of neural crest cells. *a*, Littermates (2 weeks old) with head pigmentation. *b*, Littermates (2 weeks old) with pigmented head and tail. *c*, Animal with extensive pigmentation in head, posterior trunk and hindlimbs. Note the sharp demarcation at the mid-dorsal line. *d*, Animal with two pigmented hindlimbs. Note the fine line of interspersed pigmented hair extending from the mid-dorsal line to both feet. *e*, Pigmentation in the iris. *f*, Pigmentation in the choroid as viewed from the back of the eye.

Methods. BALB/c mouse embryos were microinjected between days 8.75 and 9.25 of gestation with *in vitro*-cultured neural crest cells of C57BL/6J origin as described in the text and for Table 1.

having been deposited in the embryo proper or having migrated along abnormal pathways. (1) Pigmented cells always originated in the mid-dorsal line and when colonizing the head region, migrated to both sides mediolaterally, giving rise to symmetrical patches (see Fig. 2a–c). In all cases of distal pigmentation in the posterior trunk region, a fine line of interspersed pigmented hairs extending from the mid-dorsum down to the hind foot revealed the migratory path (see Fig. 2d). (2) Pigmentation was restricted to the head and posterior trunk region and was never seen in the anterior trunk or forelimbs, even in animals that were pigmented at the anterior as well as at the posterior end (see Fig. 2c, Table 3). (3) Both the radial pigmentation in the iris and choroid and the sharp demarcation of pigmentation along the mid-dorsal and mid-ventral lines seen in many of the chimaeras reveal the normal migratory pattern of neural crest cells and resemble the patterns observed in allophenic mice^{1,2,4,22}.

To explain the generation of chimaeric animals, it is reasonable to assume that the injection pipette has to penetrate the

Table 1 Results of injection of post-implantation embryos (BALB/c) with neural crest cells (C57BL/6J)

Age of embryos (days of gestation)	Injected	No. of embryos surviving to weaning	With pigment contributions
8.25	213	121 (57%)	0
8.75–9.25	340	212 (62%)	31
9.75	118	71 (60%)	0

BALB/c embryos were microinjected at day 8.25, between days 8.75 and 9.25, and at day 9.75 of gestation, with neural crest cells obtained from C57BL/6J neural tubes and cultured *in vitro* as described in Fig. 1 legend.

Table 2 Pigment contribution in BALB/c mice injected with C57BL/6J neural crest cells between days 8.75 and 9.25 of gestation

Age of neural crest culture	No. of animals	
	Total	With pigment
3–5 days	83	28
6–8 days	35	2
11–13 days	84	1

Injections with cells from 3–5-day-old cultures were performed at day 8 between 17:00 and 20:00 h, at around midnight on the same day, and between 06:00 and 08:00 h on day 9; no significant difference in the fraction of chimaeras obtained was observed. Most of the injections with cells from older cultures were performed at day 8.75.

Table 3 Pigment distribution in neural crest injection chimaeras

Type	No. of animals	Pigment contribution		
		Head, eyes	Anterior trunk and forelimbs	Posterior trunk (back, hindlimbs, tail)
A	19	+	—	—
B	9	—	—	+
C	3	+	—	+

yolk sac and probably the amnion, but not the embryo proper. The phenotype of animals of type C (Table 3) actually indicates that after a single injection the cells have migrated to both (opposite) ends of the animals while sparing the anterior and mid-trunk region. This suggests that the cells, after having been deposited in the amniotic cavity, are restricted as to where they can enter the embryo proper. The anterior and posterior neuropores, which close at the 18- and 30-somite stage respectively, and thus much later than the neural tube in the mid-trunk region, may serve as the points of entry. Having entered the embryo, the cells appear to have followed the normal pathways of neural crest cell migration from the mid-dorsal line mediolaterally, as indicated by the similarities of the pigment patterns seen in injection chimaeras to those of allophenic mice. If correct, this would indicate that neural crest cells, even when deposited outside their normal pathways of migration, can penetrate the epithelium and migrate to their normal locations in the epidermis to form functional pigment cells. The restriction of pigment to the mid-dorsum in many chimaeras may indicate, however, that the amelanotic host melanoblasts can interfere with the migration of the donor cells or with their 'homing-in' on the epidermis. Earlier observations have indeed demonstrated that amelanotic melanoblasts can prevent normal melanoblasts from entering the epidermis²³.

Only embryos injected between days 8.75 and 9.25 of gestation developed into animals with pigment chimaerism, whereas none of the embryos injected 12 h previously or 12 h later gave rise to pigmented mice. It is possible that the embryo at day 8.25 is physically too small to allow introduction of cells into the yolk sac and amnion. At day 9.75 of gestation, the neural tube is already closed at the anterior end, thus preventing the injected cells from entering the embryo at that site; the posterior neuropore may still be open, but technical difficulties rather than developmental constraints may interfere with successful colonization from that end at the later stage.

The manipulation of post-implantation mouse embryos by cell injection has great potential for the analysis of patterns of cell migration in early stages of mammalian morphogenesis under *in vivo* conditions. It would be particularly interesting to investigate whether other migratory cells such as primordial germ cells, which normally migrate from the allantois to the genital ridge between days 8 and 12 of gestation²⁴, could colonize an embryo after microinjection. The results described here suggest that the post-implantation embryo can be used to analyse cell interactions at critical stages of development. For example, cell surface antigens have been implicated as being important in cell adhesion and morphogenesis²⁵. Retroviral vectors²⁶ could be used to transfer cellular genes, such as those coding for cell surface antigens, into the cells prior to injection to elucidate the role of these antigens in cell migration and homing. This approach may allow us to study the cell-cell interactions in mammalian development that are crucial for morphogenesis and which, so far, have not been amenable to experimental analysis.

I thank Dr J. Price for advice in isolation and culture of neural crest cells, Dr U. Dräger for help in photography and sectioning of mouse eyes, and Dr C. Cepko for a gift of serum. This work was supported by grants HD-19105 from the NIH and PO1-CA38497 from the National Cancer Institute.

Received 1 July; accepted 25 September 1985.

1. Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **58**, 344-351 (1967).
2. Gardner, R. *Nature* **220**, 569-597 (1968).
3. Mintz, B. *A. Rev. Genet.* **8**, 411-470 (1974).
4. McLaren, A. in *Developmental and Cell Biology* Vol. 4 (Cambridge University Press, 1976).
5. Le Douarin, N. in *Developmental and Cell Biology* Vol. 12 (Cambridge University Press, 1982).
6. Le Douarin, N. *Cell* **38**, 353-360 (1984).
7. Thiery, J. *Cell Differ.* **15**, 1-15 (1984).
8. Jaenisch, R. *Cell* **19**, 181-188 (1980).
9. Stuhlmann, H., Cone, R., Mulligan, R. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7151-7155 (1984).
10. Fleischman, R. & Mintz, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5736-5740 (1979).
11. Weston, J. in *Cell Interactions and Development: Molecular Mechanisms* (ed. Yamada, K.) 153-184 (Wiley, New York, 1983).
12. Holtfreter, J. in *Epithelial-Mesenchyme Interactions* (eds Fleischmajer, P. & Billingham, R. B.) 1-30 (Williams & Wilkins, Baltimore, 1968).
13. Chibon, P. *J. Embryol. exp. Morph.* **18**, 343-358 (1967).
14. Bronner, M. & Cohen, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1843-1847 (1979).
15. Erickson, L., Tosney, K. & Weston, J. *Dev. Biol.* **77**, 142-156 (1980).
16. Ito, K. & Takeuchi, T. *J. Embryol. exp. Morph.* **84**, 49-62 (1984).
17. Sieber-Blum, M. & Cohen, A. *Dev. Biol.* **80**, 96-106 (1980).
18. Glimelius, B. & Weston, J. *Cell Differ.* **10**, 57-67 (1981).
19. Cohen, A. & Konigsberg, I. *Dev. Biol.* **46**, 262-280 (1975).
20. Deol, M. *J. Embryol. exp. Morph.* **30**, 483-489 (1973).
21. Silvers, W. *The Coat Colours of Mice* (Springer, New York, 1979).
22. Mintz, B. in *Methods in Mammalian Embryology* (ed. Daniel, J.) 186-214 (Freeman, San Francisco, 1971).
23. Mayer, T. *Dev. Biol.* **34**, 39-46 (1973).
24. Snow, M. & Monk, M. in *Current Problems in Germ Cell Differentiation* (eds McLaren, A. & Wylie, L.) 115-135 (Cambridge University Press, 1983).
25. Edelman, G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1460-1464 (1984).
26. Mulligan, R. in *Experimental Manipulation of Gene Expression*, 155-173 (Academic, New York, 1983).

Insulin rapidly stimulates tyrosine phosphorylation of a M_r -185,000 protein in intact cells

Morris F. White, Ruth Maron & C. Ronald Kahn

Research Division, Joslin Diabetes Center and Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02215, USA

Phosphotyrosine-containing proteins are minor components of normal cells^{1,2} which appear to be associated primarily with the regulation of cellular metabolism and growth^{3,4}. The insulin receptor is a tyrosine-specific protein kinase^{5,6}, and one of the earliest detectable responses to insulin binding is activation of this kinase and autophosphorylation of its β -subunit⁷⁻⁹. Tyrosine autophosphorylation activates the phosphotransferase in the β -subunit and increases its reactivity toward tyrosine phosphorylation of other substrates^{10,11}. When incubated *in vitro* with [γ -³²P]ATP and insulin, the purified insulin receptor phosphorylates various proteins on their tyrosine residues¹²⁻¹⁶. However, so far no proteins other than the insulin receptor have been identified as undergoing tyrosine phosphorylation in response to insulin in an intact cell. Here, using anti-phosphotyrosine antibodies, we have identified a novel phosphotyrosine-containing protein of relative molecular mass (M_r) 185,000 (pp185) which appears during the initial response of hepatoma cells to insulin binding. In contrast to the insulin receptor, pp185 does not adhere to wheat-germ agglutinin-agarose or bind to anti-insulin receptor antibodies. Phosphorylation of pp185 is maximal within seconds after exposure of the cells to insulin and exhibits a dose-response curve similar to that of receptor autophosphorylation, suggesting that this protein represents the endogenous substrate for the insulin receptor kinase.

To identify any phosphotyrosine-containing proteins which might appear in cells as a consequence of insulin binding, the well-differentiated and insulin-sensitive hepatoma cell line Fao^{17,18} was labelled with ³²P-orthophosphate for 2 h¹⁹. Then the cells were either treated with 100 nM insulin or left untreated and the monolayers were solubilized with Triton X-100 as described in Fig. 1 legend. Purification of the resulting extract by wheat-germ agglutinin(WGA) chromatography and immunoprecipitation with anti-insulin receptor antibody showed

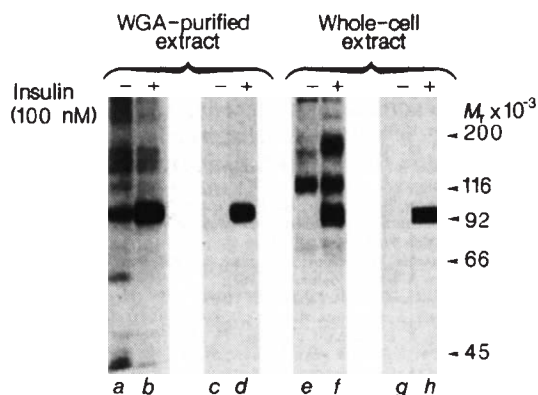


Fig. 1 Purification of phosphotyrosine-containing proteins from ^{32}P -orthophosphate-labelled Fao cells by immunoprecipitation with anti-insulin receptor and anti-phosphotyrosine antibodies. *a-d*, WGA-purified extracts immunoprecipitated with anti-receptor antibody (*a, b*) or anti-phosphotyrosine antibody (*c, d*). *e-h*, Whole-cell extracts immunoprecipitated with anti-phosphotyrosine antibody (*e, f*) or with anti-phosphotyrosine antibody, followed by anti-receptor antibody (*g, h*).

Methods. Fao cells were grown in plastic tissue culture dishes (15-cm diameter) containing 30 ml of RPMI 1640 medium (Gibco) supplemented with 10% fetal bovine serum (Gibco); 12 h before each experiment, the culture medium was changed to serum-free RPMI 1640. Confluent Fao cells were labelled for 2 h in 10 ml of phosphate-free and serum-free RPMI 1640 medium containing carrier-free ^{32}P -orthophosphate (0.5 mCi ml^{-1} ; NEN). Cells were incubated without (-) or with (+) insulin (100 nM) at 37°C for 1 min, then the experiments were stopped quickly by removing the incubation medium and freezing the cell monolayers with liquid nitrogen. The monolayers were thawed and solubilized immediately at 4°C with 2 ml of a solution containing 50 mM HEPES pH 7.4, 1% Triton X-100, 10 mM sodium pyrophosphate, 100 mM sodium fluoride, 4 mM EDTA, 2 mM sodium vanadate, 1 mg ml^{-1} aprotinin, and 2 mM phenylmethylsulphonyl fluoride. A supernatant of the whole-cell extract was prepared by scraping the cells from the dishes and sedimenting the insoluble material by centrifugation at 50,000 r.p.m. in a Beckman 70.1 Ti rotor for 60 min. WGA-purified extracts were obtained by applying the supernatant to a 0.5-cm diameter disposable column (BioRad) containing 0.2 ml of WGA-agarose (Vector). The agarose was washed with 100 ml of 50 mM HEPES pH 7.4 containing 0.1% Triton X-100, 10 mM sodium pyrophosphate, 10 mM sodium fluoride, 4 mM EDTA and 2 mM sodium vanadate, and the bound glycoproteins were eluted with two 0.5-ml portions of this wash solution containing 300 mM *N*-acetylglucosamine (Sigma). Specific immunoprecipitates were obtained from WGA-purified extracts (*a-d*) or from whole-cell extracts (*e-h*) using anti-phosphotyrosine antibodies prepared according to the method of Pang *et al.*³⁴, or anti-insulin receptor antibodies (B2) obtained from patients³¹. After incubation with the extract at 4°C for 2 h, the antibodies were immobilized on Pansorbin (Calbiochem) and the precipitates were washed three times with a solution containing 50 mM HEPES, 1% Triton X-100 and 0.1% SDS. Proteins were eluted from the phosphotyrosine-antibody complex by addition of 10 mM *p*-nitrophenylphosphate to the wash solution⁹ and from the anti-receptor antibody with SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer³¹. The eluted proteins were reduced with 100 mM dithiothreitol (DTT; BioRad) and the phosphoproteins were separated by SDS-PAGE on 7.5% resolving polyacrylamide gels as described previously³². The phosphoproteins were identified by autoradiography (6 h exposure) of the stained and dried gels using Kodak X-Omat film and an intensifying screen. The WGA-purified extract was immunoprecipitated with anti-receptor antibody (*a, b*) or anti-phosphotyrosine antibody (*c, d*). The whole-cell extract was immunoprecipitated with anti-phosphotyrosine antibody and the precipitated phosphoproteins were either separated by PAGE (*e, f*) or immunoprecipitated a second time with anti-receptor antibodies before SDS-PAGE (*g, h*).

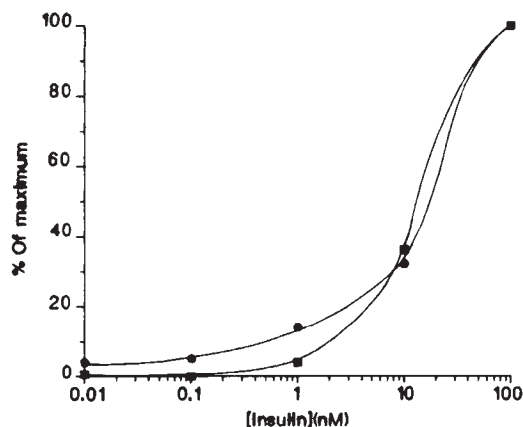


Fig. 2 Dose-response curve for effect of insulin on tyrosine phosphorylation. Labelled Fao cells were incubated in the presence of insulin for 10 min. Phosphoproteins were immunoprecipitated from the whole-cell extracts with the anti-phosphotyrosine antibody, eluted, separated by SDS-PAGE and identified by autoradiography as described in Fig. 1 legend. The intensities of pp185 (●) and the β -subunit (■) were measured by scanning densitometry and are presented as a percentage of the maximum stimulation. The maximum values, measured at 100 nM insulin, were 603 and 1,680 for pp185 and the β -subunit, respectively. The background measured in the absence of insulin has been subtracted.

a single insulin-stimulated phosphoprotein of M_r 95,000. This band was present in the basal state and increased about fourfold during insulin stimulation for 1 min (Fig. 1*a, b*). We have shown previously that this band corresponds to the β -subunit of the insulin receptor which undergoes tyrosine phosphorylation in response to insulin stimulation^{7,9,19}. When the WGA-purified extract was immunoprecipitated with a polyclonal anti-phosphotyrosine antibody a similar result was observed, except that no phosphorylation of the β -subunit was seen under basal conditions because in the intact cells the β -subunit does not contain phosphotyrosine before insulin stimulation (Fig. 1*c, d*)^{9,19}. During 1 h of incubation with insulin, the amount of ^{32}P -orthophosphate incorporated into the proteins eluted from the WGA-agarose was constant, suggesting that insulin stimulation of receptor phosphorylation does not result from a general increase in the specific activity of cellular ATP.

In an attempt to identify other phosphotyrosine-containing proteins, Fao cells were labelled and subjected to immunoprecipitation with anti-phosphotyrosine antibodies before WGA chromatography. In the absence of insulin, a single major phosphoprotein of M_r 120,000 was immunoprecipitated from the whole-cell extract (Fig. 1*e*). After incubation for 1 min with 100 nM insulin there was no change in pp120 (Fig. 1*f*), but we observed two new phosphoproteins of M_r 95,000 and 185,000 (Fig. 1*f*). pp95 is the β -subunit of the insulin receptor and was quantitatively immunoprecipitated by anti-insulin receptor serum (B2) (Fig. 1*g, h*). pp185, on the other hand was not recognized by this antibody, suggesting that it is a novel insulin-stimulated phosphotyrosine-containing protein that is antigenically distinct from the insulin receptor (Fig. 1*f, h*). The insulin dose-response curve for phosphorylation of pp185 was nearly identical to that for the β -subunit, supporting the notion that pp185 is a substrate for tyrosine phosphorylation by the insulin receptor kinase (Fig. 2).

The phosphoamino-acid composition of pp185, pp120 and the β -subunit of the insulin receptor was determined by partial acid hydrolysis and separation of the amino acids by high-voltage electrophoresis¹⁹. After insulin stimulation, each protein contained about equal amounts of phosphotyrosine and phosphoserine (Fig. 3); pp185 also contained substantial amounts

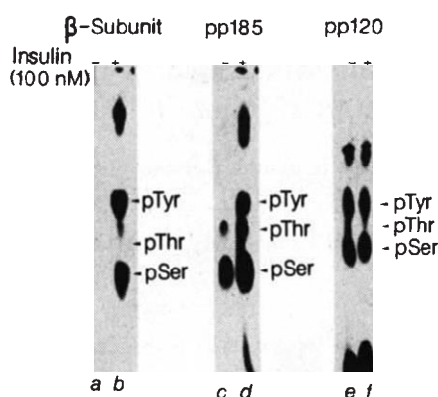


Fig. 3 Phosphoamino-acid analysis of pp185, pp120 and the β -subunit of the insulin receptor immunoprecipitated from whole-cell extracts with the anti-phosphotyrosine antibody. Fao cells were incubated in the absence (–) or presence (+) of insulin (100 nM) for 1 min. Phosphoproteins were immunoprecipitated from the whole-cell extracts with the anti-phosphotyrosine antibody and specifically eluted with *p*-nitrophenylphosphate, reduced with DTT and separated by SDS-PAGE. Dried gel fragments containing the β -subunit (a, b), pp185 (c, d) and pp120 (e, f) were incubated with trypsin (50 μ g ml^{–1}) as described previously^{9,31}. The phosphoamino-acid composition of the eluted tryptic peptides was determined as described previously³¹. The phosphoamino-acid standards were visualized by reaction with ninhydrin and the radioactive amino acids were detected by autoradiography for 24 h.

of phosphothreonine. Although there was no distinct evidence of pp185 before insulin stimulation, there was a faint band that migrated slightly faster than pp185 which contained a small amount of phosphotyrosine (Figs 1e, 3c). Assuming that this is the same protein, insulin stimulated its phosphotyrosine content at least 10-fold (Fig. 3d). A significant increase in the amount of phosphoserine and phosphothreonine was also observed in pp185 after insulin stimulation, but whether this arises from *de novo* phosphorylation or tyrosine phosphorylation of previously phosphorylated pp185 is unknown. Before and after insulin stimulation, pp120 contained nearly equal amounts of phosphotyrosine, suggesting that it may be a substrate for an unidentified and constitutively activated tyrosine kinase in Fao cells.

Phosphorylation of the β -subunit is one of the earliest molecular responses that occurs in cells following insulin binding^{9,19}. As we have reported previously¹⁹, within 20 s after insulin stimulation (100 nM), phosphorylation of the β -subunit is maximal and remains at this elevated level for at least 1 h during uninterrupted exposure to insulin (Fig. 4). Phosphorylation of pp185 was also maximal by 30 s after exposure of the Fao cells to insulin, but it decreased during 60 min of continued insulin stimulation (Fig. 4); this decrease was not the result of a change in the specific activity of cellular [γ -³²P]ATP as there was no change in the level of phosphorylation of pp120 or of the β -subunit, which is rapidly dephosphorylated when the labelled ATP pool is diluted with a chase of unlabelled phosphate²⁰. Phosphoamino-acid analysis of pp185 indicated that concentrations of phosphoserine, phosphothreonine and phosphotyrosine decreased in parallel during this time interval (data not shown). These results suggest that following phosphorylation of pp185 there may be activation of a phosphoprotein phosphatase that recognizes pp185, but not the insulin receptor or pp120, as a substrate for dephosphorylation. These findings are consistent with the notion that the action of insulin may be regulated in the intact cell by both phosphorylation and dephosphorylation reactions²¹.

In the experiment shown in Fig. 4, phosphoproteins of M_r 75,000 and >200,000 were also detected by immunoprecipitation with the anti-phosphotyrosine antibody. The band of higher relative molecular mass showed similar kinetics to pp185,

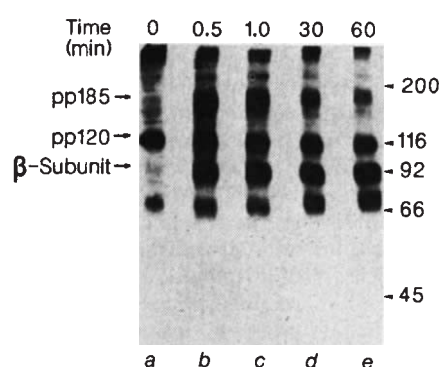


Fig. 4 The time course of insulin-stimulated tyrosine phosphorylation. Fao cells were incubated in the absence (0 min) or presence of insulin for the indicated time intervals. Phosphoproteins were immunoprecipitated from the whole-cell extracts with the anti-phosphotyrosine antibody, eluted with *p*-nitrophenylphosphate and identified by SDS-PAGE and autoradiography as described in Fig. 1 legend.

whereas pp75 was present in the basal state and tended to increase after 30–60 min of insulin stimulation. These phosphoproteins were less abundant than the others described above and were observed only after prolonged exposure of the autoradiograms. Phosphoproteins with M_r s of 300,000–200,000 were detected by anti-phosphotyrosine antibodies in human fibroblasts stimulated with platelet-derived growth factor (PDGF)²².

Like the epidermal growth factor (EGF) receptor, PDGF receptor and several oncogene products, the insulin receptor is a tyrosine-specific protein kinase. Previous studies have successfully identified several cellular substrates (pp34, pp41, pp81, enolase, lactate dehydrogenase, phosphoglycerate mutase, vinculin) for the EGF receptor^{23–26}, PDGF receptor^{22,27} and pp60^{src} (refs 23, 24, 26, 28), but so far no endogenous substrates for the insulin receptor have been identified. The present study indicates that in addition to stimulating autophosphorylation of the insulin receptor, insulin stimulates tyrosine phosphorylation of a protein in intact hepatoma cells that has a M_r of 185,000. This protein (pp185) may have gone undetected by previous studies for several reasons. First, its rapid phosphorylation and dephosphorylation in insulin-stimulated cells mean that it can be detected only at early time points. Second, WGA chromatography is often used to purify cell extracts before studying insulin receptor phosphorylation²⁹, but as pp185 does not bind to this lectin it would not be detected in such experiments. Third, sodium vanadate, an inhibitor of dephosphorylation^{19,30}, was used in our studies during cell lysis and protein purification, but it has not been used routinely in the past. Finally, two-dimensional gel electrophoresis, which has been successful in identifying many tyrosine kinase substrates^{23,28}, may miss phosphoproteins with *pI* values <5.5 (ref. 27) or high- M_r phosphoproteins due to the background phosphorylation that occurs in this region²². These limitations can be avoided by using anti-phosphotyrosine antibodies, which provides specific probes for studying tyrosine phosphorylation events that occur during the initial response of cells to insulin.

Although the exact nature of pp185 is unknown, we can eliminate certain possibilities. pp185 does not bind to WGA and was not immunoprecipitated by the anti-insulin receptor serum B2, which recognized the mature receptor and the glycosylated and non-glycosylated (H. A. Hedo, personal communication) forms (M_r 155,000–190,000) of the precursor^{31,32}. Furthermore, the tryptic phosphopeptide map of pp185 does not correspond to the profile of β -subunit (data not shown). Thus, pp185 is not the insulin receptor. Fao cells show no specific EGF binding and no detectable tyrosine phosphorylation during EGF stimulation (data not shown), suggesting that pp185 is not related to

the EGF receptor (M_r 175,000). pp185 is also probably not the PDGF receptor (M_r 185,000) as the latter binds to WGA^{26,33}.

Immunoprecipitation of ³⁵S-methionine-labelled Fao cells by the anti-phosphotyrosine antibody did not reveal any band at M_r 185,000, although both the α - and β -subunits were easily detected (data not shown). Thus, pp185 is possibly a minor protein that is recognized with high affinity by the insulin receptor. Whether pp185 is directly phosphorylated by the insulin receptor kinase or is the result of activation of another tyrosine kinase in response to insulin is not yet known.

The only other prominent phosphoprotein in Fao cells that is recognized by the anti-phosphotyrosine antibody has a M_r of 120,000 and is constitutively phosphorylated on tyrosine residues. The nature of this phosphoprotein is unknown. Others have reported that ATP citrate lyase (M_r 116,000) is equally immunoprecipitated with a monoclonal antibody against phenylphosphonate before and after EGF stimulation of A431 cells because it contains phosphohistidine, which cross-reacts with this antibody²⁵; however, phosphoamino-acid analysis confirmed the presence of phosphotyrosine in pp120. Insulin stimulates tyrosine phosphorylation of a M_r -120,000 protein in a Triton X-100 extract of rat hepatocyte plasma membranes purified by WGA chromatography²⁹. It is unknown whether pp120 bears any similarity to the hepatocyte protein but they are probably different as pp120 identified here does not bind to WGA.

Our results provide the first demonstration, in an intact cell, of insulin-stimulated tyrosine phosphorylation of a protein other than the insulin receptor. Identification of pp185 should elucidate the relationship between the insulin receptor kinase and the mechanism of insulin action.

This work was supported in part by grants AM31036 and AM29770 to C.R.K. and a fellowship (AM0716301) to M.F.W. from the Institute of Health and Human Development, NIH, USPHS. We thank Drs D. T. Pang, B. R. Sharma and J. A. Shafer for the gift of phosphotyrosine binding antibody and for making available prior to publication the procedures for preparation of the antibody.

Received 22 July; accepted 18 September 1985.

- Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
- Sefton, B. M., Hunter, T., Beemon, K. & Eckhart, W. *Cell* **20**, 807-816 (1980).
- Bishop, M. A. *Rev. Biochem. Sci.* **32**, 301-354 (1983).
- Heldin, C.-H. & Westermark, B. *Cell* **37**, 9-20 (1984).
- Kasuga, M., Fujita-Yamaguchi, Y., Bliethe, D. L. & Kahn, C. R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2137-2141 (1980).
- Petrizzelli, L., Herrera, R. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3327-3331 (1984).
- Kasuga, M., Karlsson, F. A. & Kahn, C. R. *Science* **215**, 185-187 (1982).
- Kasuga, M., Zick, Y., Bliethe, D. L., Crettaz, M. & Kahn, C. R. *Nature* **298**, 667-669 (1982).
- Pang, D. T., Sharma, B. R., Shafer, J. A., White, M. F. & Kahn, C. R. *J. biol. Chem.* **260**, 7131-7136 (1985).
- Rosen, O. M., Herrera, R., Olowe, Y., Petrizzelli, L. M. & Cobb, M. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3237-3240 (1983).
- Yu, K.-T. & Czech, M. P. *J. biol. Chem.* **259**, 5277-5286 (1984).
- Kasuga, M., Fujita-Yamaguchi, Y., Bliethe, D. L., White, M. F. & Kahn, C. R. *J. biol. Chem.* **258**, 10973-10979 (1983).
- Stadtmauer, L. A. & Rosen, O. M. *J. biol. Chem.* **258**, 6682-6685 (1983).
- Zick, Y. *et al. Eur. J. Biochem.* **137**, 631-637 (1983).
- Haring, H. U. *et al. J. Cell Biochem.* **27**, 171-182 (1985).
- White, M. F. & Sale, E. M. *Diabetes* **33**(Suppl. 1), 31A (1984).
- Deschatrete, J., Moore, E. E., Dubois, M., Cassio, D. & Weiss, M. C. *Somat. Cell Genet.* **5**, 697-718 (1979).
- Deschatrete, J. & Weiss, M. C. *Biochimie* **11**, 1603-1611 (1974).
- White, M. F., Takayama, S. & Kahn, C. R. *J. biol. Chem.* **260**, 9470-9478 (1985).
- Haring, H. U., Kasuga, M., White, M. F., Crettaz, M. & Kahn, C. R. *Biochemistry* **23**, 3298-3306 (1984).
- Denton, R. M., Brownsey, R. W. & Belshan, G. J. *Diabetologia* **21**, 347-362 (1981).
- Ek, B. & Heldin, C.-H. *J. biol. Chem.* **259**, 11145-11152 (1984).
- Cooper, J. A. & Hunter, T. *J. Cell Biol.* **91**, 878-893 (1983).
- Erikson, E., Shealy, D. J. & Erikson, R. L. *J. biol. Chem.* **256**, 11381-11384 (1981).
- Frackelton, A. R., Ross, A. H. & Eisen, H. N. *Molec. cell Biol.* **3**, 1343-1352 (1983).
- Sefton, B. M. & Hunter, T. *Cell* **24**, 165-174 (1981).
- Frackelton, A. R., Tremble, P. M. & Williams, L. T. *J. biol. Chem.* **259**, 7909-7915 (1984).
- Cooper, J. A., Reiss, N. A., Schwartz, R. J. & Hunter, T. *Nature* **302**, 218-223 (1983).
- Rees-Jones, R. & Taylor, S. J. *J. biol. Chem.* **260**, 4461-4467 (1985).
- Swarup, G., Cohen, S. & Garbers, D. L. *Biochem. biophys. Res. Commun.* **113**, 80-86 (1982).
- Kasuga, M., White, M. F. & Kahn, C. R. *Meth. Enzym.* **109**, 609-621 (1985).
- Ronnett, G. V., Knutson, V. P., Kohanski, R. A., Simpson, T. L. & Lane, M. D. *J. biol. Chem.* **259**, 4566-4575 (1984).
- Daniel, T. O., Tremble, P. M., Frackelton, A. R. & Williams, L. T. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2684-2687 (1985).
- Pang, D. T., Sharma, B. R. & Shafer, J. A. *Archs. Biochem. Biophys.* **242**, 176-186 (1985).

Aspirin causes short-lived inhibition of bradykinin-stimulated prostacyclin production in man

Dennis J. Heavey, Susan E. Barrow, Nicola E. Hickling & James M. Ritter

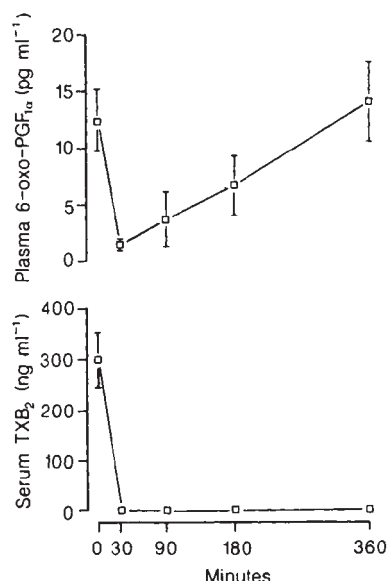
Department of Clinical Pharmacology, Royal Postgraduate Medical School, London W12 0HS, UK

Acetylsalicylic acid (aspirin) inhibits prostanoid synthesis^{1,2} by irreversible acetylation of fatty acid cyclooxygenase (EC 1.14.99.1)³. It thereby inhibits synthesis of pro-aggregatory thromboxane A₂ (TXA₂) by platelets^{4,5} and is widely used in the treatment and prophylaxis of vascular disease. Its efficacy, however, may be reduced since it also inhibits formation of prostacyclin (PGI₂)^{5,6} which is a vasodilator and anti-aggregatory agent^{7,8}. There is uncertainty over the optimum dose regimen for aspirin since although it inhibits platelet thromboxane production for many days⁴, the magnitude and duration of its effect on PGI₂ production by vascular endothelium *in vivo* is unknown. Resting plasma concentrations of PGI₂ (measured as the stable hydrolysis product 6-oxo-PGF_{1 α}) are at or below the limit of sensitivity of the most sensitive assays⁹ and cannot therefore be used to demonstrate a reduction in production. Bradykinin stimulates PGI₂ synthesis by cultured human vascular endothelial cells¹⁰ and we have shown that it stimulates PGI₂ production by man *in vivo*¹¹. We report here that an oral dose of aspirin (600 mg) causes rapid and substantial inhibition of bradykinin-stimulated PGI₂ production, but recovery occurs within 6 hours; this implies that endothelial PGI₂ synthesis would be spared most of the time during dosing once daily with even this relatively large dose of aspirin.

Seven healthy men aged 21-39 years, weight 68.5-85 kg, were studied. The protocol was approved by the Ethics Committee of the Royal Postgraduate Medical School, Hammersmith Hospital and subjects gave written informed consent. Bradykinin (RIA Biochemicals) was obtained as a solid, and a sterile, pyrogen-free solution in water was prepared, lyophilized and stored at -80 °C. On the day of each study bradykinin was dissolved in sterile isotonic saline and passed through a 0.22- μ m filter (Millipore). PGI₂ production was stimulated by short intravenous infusions via an indwelling cannula in one forearm. Blood was sampled from a cannula in the other forearm. Bradykinin causes vasodilatation and swelling around the eyes and lips with a feeling of general discomfort, and the dose used, although the same within each subject, was varied from subject to subject depending on individual tolerance. The dose ranged from 0.15 to 0.8 μ g per kg body weight per min given as a single 5-min infusion or two consecutive 5-min infusions, the second at twice the dose of the first. The latter protocol improved tolerance in some subjects. During bradykinin infusions, blood pressure was recorded at 1-min intervals (Arteriosonde, Roche) and the electrocardiogram and skin temperature recorded continuously. Infusions were given before, and 30, 90, 180 and 360 min after drinking 600 mg soluble aspirin (Bayer) in 50 ml water. Blood samples (~20 ml) for measurement of the 6-oxo-PGF_{1 α} concentration were collected before and at the end of each bradykinin infusion. The ability of platelets to produce TXA₂ was determined by measurement of the stable hydrolysis product, TXB₂, in serum. Blood (2 ml) was taken immediately before each bradykinin infusion and incubated in a plain glass tube at 37 °C for 1 h⁴. TXB₂ and 6-oxo-PGF_{1 α} were measured by a highly specific and sensitive method based on gas chromatography/negative-ion chemical ionization mass spectrometry (GC/NICIMS)^{9,12}.

Plasma concentrations of 6-oxo-PGF_{1 α} are shown in Table 1. Concentrations before bradykinin infusion were usually close to, or below, the limit of sensitivity of the assay (generally 1 pg ml⁻¹). In a few cases, especially in the control period, resting concentrations were higher, probably as a result of

Fig. 1 Effect of aspirin on serum TXB₂ concentration and bradykinin-stimulated PGI₂ production. Values shown are means $\pm$ s.e. ($n=7$). Plasma 6-oxo-PGF_{1 α} concentrations are those measured at the end of each bradykinin infusion (Table 1). Blood (2 ml) taken immediately before each bradykinin infusion was allowed to clot in a plain glass tube at 37 °C for 1 h and serum TXB₂, the stable hydrolysis product of TXA₂, measured by GC/NICIMS. The serum TXB₂ concentration reflects the ability of platelets to synthesize TXA₂. In common with other investigators⁴, and in contrast to the effect on bradykinin-stimulated PGI₂ production, serum TXB₂ remained more than 99% inhibited 260 min after aspirin.



endothelial trauma during cannulation^{13,14}. In each subject the plasma 6-oxo-PGF_{1 α} concentration increased during the control bradykinin infusion (before aspirin), although 2 subjects (6 and 7) who tolerated only the lowest doses of bradykinin showed little stimulation. PGI₂ production was substantially inhibited 30 min after aspirin administration. In 5 of 7 subjects the stimulated 6-oxo-PGF_{1 α} concentration was below the detection limit. Assuming that the concentration in those cases was actually at the limit of detection, the mean inhibition of stimulated 6-oxo-PGF_{1 α} concentration was 86%. Recovery had usually occurred by 6 h. The serum TXB₂ concentration was inhibited by more than 99% at all times after aspirin. PGI₂ is a more potent vasodilator than bradykinin but even at 30 min after aspirin there was no attenuation of the vasodilator effects of bradykinin (data not shown). Figure 1 compares the effect of aspirin on TXA₂ and PGI₂ production. To rule out the possibility that the reduced PGI₂ synthesis at 30 min was a result of desensitization to bradykinin rather than inhibition by aspirin, three of the subjects were studied using the same protocol with placebo instead of aspirin. There was no effect on either serum TXA₂ generation or bradykinin-stimulated PGI₂ production (Fig. 2). The rapid recovery of PGI₂ after aspirin led us to believe that a dose of aspirin once daily would not cause a cumulative effect on bradykinin-stimulated PGI₂ synthesis and one subject was therefore given 600 mg daily for 7 days. On day 8, 18 h after

the last dose of aspirin, bradykinin infusion produced an increase in the plasma 6-oxo-PGF_{1 α} concentration which was again inhibited 30 min after a further dose of aspirin (600 mg) and recovered within hours (Fig. 3).

The tissue of origin of bradykinin-stimulated PGI₂ is likely to be vascular endothelium. Cultured human endothelial cells synthesize PGI₂ in response to bradykinin¹⁰, and we have found that during constant infusion of bradykinin the plasma concentration of 6-oxo-PGF_{1 α} rises to a plateau within 5 min (data not shown). Czervionke found that aspirin causes over 90% inhibition of PGI₂ production by cultured human endothelial cells, and recovery by synthesis of new cyclooxygenase is apparent within 2 h of washing¹⁵. Jaffe and Weksler obtained similar results with substantial recovery by 24 h¹⁶. The rate of recovery of cyclooxygenase activity *in vitro* after aspirin is highly dependent on culture conditions¹⁷, emphasizing the need for *in vivo* studies. Following an oral dose of 600 mg soluble aspirin, the peak plasma acetylsalicylic acid concentration occurs at 15 min, with very rapid elimination (elimination $t_{1/2}$ = 0.31 h)¹⁸. Tissues are therefore exposed to acetylsalicylic acid for only a short time. Our finding of a rapid inhibition of PGI₂ production with recovery in hours, although consistent with *in vitro* studies of endothelial cells, is seemingly at variance with the only previous *in vivo* study in man¹⁹, and with studies on human vascular tissue *ex vivo*²⁰⁻²². FitzGerald found that aspirin causes

Table 1 Effect of aspirin on bradykinin(BK)-stimulated production of 6-oxo-PGF_{1 α} (pg ml⁻¹)

Subject	Time after aspirin									
	Control		30		90		180		360	
	Pre	BK	Pre	BK	Pre	BK	Pre	BK	Pre	BK
1	1.7	12.9	<1	4.0	<1	17.1	<1	18.9	<1	26.9
2	1.1	10.2	2.2	<1	1.5	NA	<1	<1	<1	4.0
3	7.3	26.2	1.2	1.2	<1	<1	<1	3.6	<1	13.8
4	4.6	13.7	<1	<1.2	<1	1.1	<1	8.0	<1	22.0
5	<1	14.4	<1	<1	<1	<1	<1	13.0	<1	19.9
6	3.4	5.1	<1	<1	<1	<1	<1	<1	<1	8.3
7	<1	4.5	<1	<1	<1	<1	<1	<1	<1	3.3

PGI₂ production was stimulated by short intravenous infusions of bradykinin (BK). The dose of bradykinin was limited by its cardiovascular effects and adjusted for each subject on the basis of preliminary studies. Infusions were given before and at 30, 90, 180 and 360 min after a dose of soluble aspirin (600 mg in 50 ml water). Plasma concentrations of 6-oxo-PGF_{1 α} , stable hydrolysis product of PGI₂, were measured before and at the end of each bradykinin infusion by GC/NICIMS⁹. The limit of sensitivity of the assay is generally 1 pg ml⁻¹ for a 20 ml blood sample. Concentrations before bradykinin infusions were usually at or below the limit of detection. In each case there was an increase in the plasma 6-oxo-PGF_{1 α} concentration during the initial bradykinin infusion, although in subjects 6 and 7, who tolerated only the lowest doses of bradykinin, the rise was small. The increase in plasma 6-oxo-PGF_{1 α} concentration is inhibited 30 min after aspirin ($2\alpha = 0.02$, Wilcoxon's signed rank test) and recovery is seen over the following 360 min. NA, result not available.

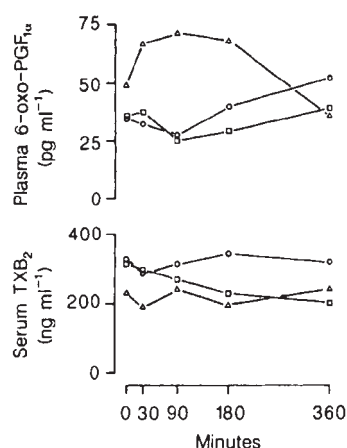


Fig. 2 Effect of aspirin placebo on serum TXB₂ concentration and bradykinin-stimulated prostacyclin production in three subjects. There was no inhibition of bradykinin-stimulated prostacyclin production, showing that the inhibition seen following aspirin is not a result of desensitization caused by the short intervals between bradykinin infusions.

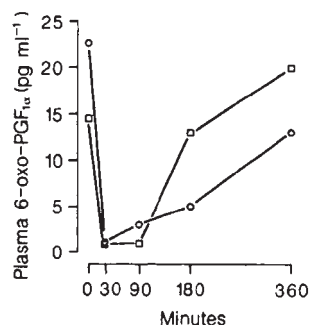


Fig. 3 Effect of chronic aspirin administration on bradykinin-stimulated PGI₂ production. One subject was studied twice. On one occasion he received 600 mg aspirin, having been drug-free for the previous 2 weeks (□), and on the second occasion he had taken aspirin 600 mg once daily for the previous 7 days, the last dose being 18 h before the dose taken on the study day (○). There was no cumulative effect of aspirin administered once daily on bradykinin-stimulated PGI₂ production, the time course of the effect being similar in the two studies. On day 8 of the chronic study, serum TXB₂ was >99% inhibited in the control (pre-aspirin) period compared with the same period of the acute study (data not shown).

only a partial (75%) inhibition of excretion of a urinary metabolite of PGI₂ even at doses of 2,600 mg daily, and recovery took more than 3 days¹⁹. The urinary metabolite, however, reflects production from many tissues with contributions from non-endothelial sources including vascular smooth muscle^{23,24}, pericardium²⁵, and gastric mucosa²⁶. If these tissues re-synthesize cyclooxygenase more slowly than does vascular endothelium, this could account for delayed recovery of the urinary PGI₂ metabolite following aspirin. If vascular endothelium contributes substantially to the urinary metabolite and recovers within hours of aspirin administration, then a 24-h urine collection will show only partial inhibition of excretion. Human vascular tissue removed 24–48 h after aspirin and incubated *ex vivo* still shows inhibition of PGI₂ production²². Much of the PGI₂ synthesized by such fragments, however, is derived from non-endothelial tissues²³, which may have a slower rate of re-synthesis of cyclooxygenase. Weksler has found that the endothelial surface of human saphenous veins has largely recovered PGI₂ synthesizing capacity by 24 h after a dose of

aspirin²⁷. Rapid endothelial cyclooxygenase synthesis could explain the finding that bovine aortic endothelial cyclooxygenase concentration is approximately 20-fold greater than that in the media²⁸.

Our results have therapeutic implications. There has been much effort expended in attempting to determine a dose of aspirin that is sufficiently low to cause selective inhibition of platelet TXA₂ synthesis without affecting PGI₂ production by vascular endothelium. The dose used in the present study is, by these standards, large. Our results indicate, however, that even this large dose has only a short-lived effect on endothelial PGI₂ production. More important than the use of a very low dose of aspirin is the need for a long dose interval, perhaps as long as several days²⁹ and certainly not less than 6 h.

This work was supported in part by an MRC Programme Grant and by the British Heart Foundation and University of London. D.J.H. is an MRC Training Fellow.

Received 4 July; accepted 2 October 1985.

1. Ferreira, S. H., Moncada, S. & Vane, J. R. *Nature new Biol.* **231**, 237–239 (1971).
2. Smith, J. B. & Willis, A. L. *Nature new Biol.* **231**, 235–237 (1971).
3. Roth, G. J. & Majerus, P. W. *J. clin. Invest.* **56**, 624–632 (1975).
4. Patrono, C. *et al. Thromb. Res.* **17**, 317–327 (1980).
5. Marcus, A. J. *New Engl. J. Med.* **297**, 1284–1285 (1977).
6. Burch, J. W., Baenziger, N. L., Stanford, N. & Majerus, P. W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5181–5184 (1978).
7. Moncada, S., Gryglewski, R., Bunting, S. & Vane, J. R. *Nature* **263**, 663–665 (1976).
8. Gryglewski, R., Bunting, S., Moncada, S., Flower, R. J. & Vane, J. R. *Prostaglandins* **12**, 685–713 (1976).
9. Blair, I. A., Barrow, S. E., Waddell, K. A., Lewis, P. J. & Dollery, C. T. *Prostaglandins* **23**, 579–589 (1982).
10. Hong, S. L. *Thromb. Res.* **18**, 789–795 (1980).
11. Barrow, S. E. *et al. Br. J. Pharmac.* (in the press).
12. Waddell, K. A. *et al. Int. J. Mass Spectrom. ion Phys.* **48**, 233–236 (1983).
13. Ritter, J. M., Barrow, S. E., Blair, I. A. & Dollery, C. T. *Lancet* **i**, 317–319 (1983).
14. Dollery, C. T. *et al. Atherosclerosis: Mechanisms and Approaches to Therapy* 105–123 (ed. Miller, N. E.) (Raven, New York, 1983).
15. Czervionke, R. L., Smith, J. B., Fry, G. L., Hoak, J. C. & Haycraft, D. L. *J. clin. Invest.* **63**, 1089–1092 (1979).
16. Jaffe, E. A. & Weksler, B. B. *J. clin. Invest.* **63**, 532–535 (1979).
17. Dejana, E. *et al. Biochem. Pharmac.* **32**, 710–713 (1983).
18. Seymour, R. A. & Rawlins, M. D. *Br. J. clin. Pharmac.* **13**, 807–810 (1982).
19. FitzGerald, G. A. *et al. J. clin. Invest.* **71**, 676–688 (1983).
20. Weksler, B. B. *et al. New Engl. J. Med.* **308**, 800–805 (1983).
21. Preston, F. E. *et al. New Engl. J. Med.* **304**, 76–79 (1981).
22. Hanley, S. P., Cockbill, S. R., Bevan, J. & Heptinstall, S. *Lancet* **i**, 969–971 (1981).
23. Moncada, S., Herman, A. G., Higgs, E. A. & Vane, J. R. *Thromb. Res.* **11**, 323–344 (1977).
24. Baenziger, N. L., Becherer, P. R. & Majerus, P. W. *Cell* **16**, 967–974 (1979).
25. Dusting, G. J. & Nolan, R. D. *Br. J. Pharmac.* **74**, 553–562 (1981).
26. Whittle, B. J. R., Higgs, G. A., Eakins, K. E., Moncada, S. & Vane, J. R. *Nature* **284**, 271–273 (1980).
27. Weksler, B. B., Tack-Goldman, K., Subramanian, V. A. & Gay, W. A. *Circulation* **71**, 332–340 (1985).
28. DeWitt, D. L., Day, J. S., Sonnenburg, W. K. & Smith, W. L. *J. clin. Invest.* **72**, 1882–1888 (1983).
29. Moncada, S. *Br. J. Pharmac.* **76**, 3–31 (1982).

Expression of N-myc in teratocarcinoma stem cells and mouse embryos

Aya Jakobovits*, Manfred Schwab†, J. Michael Bishop† & Gail R. Martin*

* Department of Anatomy, and † G. W. Hooper Foundation and Department of Microbiology and Immunology, University of California, School of Medicine, San Francisco, California 94143, USA

The N-myc gene, which is distantly related to the proto-oncogene c-myc, was first detected as an amplified sequence in human neuroblastoma cell lines and tumours^{1,2}. It has since been revealed that there is up to a 300-fold amplification of N-myc DNA in almost 50% of advanced metastatic human neuroblastomas, whereas amplification is not detected in less advanced tumours that have a better prognosis (ref. 3 and M.S., unpublished data). Although expression of N-myc is detectable in all neuroblastoma cell lines and tumours examined, its level is greatly enhanced when the N-myc gene is amplified^{4,5}. Recently, it has been shown that on co-transfection with the c-Ha-ras (EJ) gene, N-myc can induce

the malignant transformation of rat embryo fibroblasts⁶. Taken together, these data imply a function for *N-myc* in the development and/or progression of human neuroblastomas. Surveys indicate that *N-myc* also may be amplified and/or expressed in two other types of human tumours and cell lines derived from them: retinoblastomas and small cell lung cancers^{4,5,7,8}. Here, we report that *N-myc* is expressed at high levels in mouse and human teratocarcinoma stem cells, thus identifying another tumour cell type that expresses the *N-myc* gene. In addition, we found that *N-myc* is abundantly expressed in mouse embryos at mid-gestation and that its expression appears to decrease as the embryo approaches term. In the adult mouse, *N-myc* is expressed at an approximately fivefold lower level in the brain than in teratocarcinoma stem cells and embryos, and at even lower levels in the adult testis and kidney. Our data represent the first demonstration of expression of the *N-myc* gene in normal cells, and suggest that *N-myc* may be involved in mammalian embryogenesis.

Two mouse teratocarcinoma stem cell lines (embryonal carcinoma (EC) cells), PSA-1 (refs 9, 10) and F9 (ref. 11), were tested for the expression of *N-myc*. Both cell lines were isolated from embryo-derived teratocarcinomas and form tumours when injected subcutaneously (s.c.) into mice. The two cell lines are closely related to normal early embryonic cells¹² but differ in their ability to differentiate. Whereas PSA-1 cells are pluripotent and on differentiation both *in vivo* and *in vitro* give rise to a great variety of differentiated tissues, F9 cells have a very limited differentiative capacity. *N-myc* expression in PSA-1 and F9 cells was evaluated by Northern blot analysis of poly(A)⁺ cytoplasmic RNA isolated from the EC cells using a human genomic *N-myc* clone, pNb-1 (ref. 1). Both mouse EC cell lines were found to express a 3.2-kilobase (kb) *N-myc* transcript (Fig. 1). Although this transcript appeared to be relatively abundant as it was readily detected in 1 µg of total cytoplasmic RNA, it was at least 10-fold less abundant than the *N-myc* transcript of similar size observed in human neuroblastoma cells (Kelly line) in which the *N-myc* sequence is amplified ~100-fold¹ (data not shown). *N-myc* expression, equivalent to that found in PSA-1 cells, was also detected in poly(A)⁺ cytoplasmic RNA from cells of an embryonic stem cell (ESC) line (Fig. 1). Such ES cells, which are isolated from the fetal portion (inner cell mass, ICM) of a normal 4.5-day-old mouse embryo, are indistinguishable from tumour-derived teratocarcinoma stem cells and generally form teratocarcinomas with a wide variety of differentiated cell types when injected s.c. into mice^{13,14}. Thus, three tumorigenic cell lines which are known to have many properties in common with, or are directly derived from, the ICM of a normal mouse embryo, abundantly express *N-myc*. We also found that the cell line NTera2 clone D1 (ref. 15), derived from a human teratocarcinoma, abundantly expresses the 3.2-kb *N-myc* transcript (data not shown), whereas a mouse embryo fibroblast cell line, STO⁹, and a teratocarcinoma-derived endodermal cell line, PYS¹⁶, which is not tumorigenic, show no detectable *N-myc* expression (Fig. 1).

To determine whether this abundant expression in mouse EC and ES cells is due to *N-myc* gene amplification, we performed a Southern blot analysis of cellular genomic DNA (see Fig. 2). PSA-1, ES and F9 cells, in which *N-myc* is expressed at relatively high levels, appear to have the same number of copies of the gene as do mouse neuroblastoma cells, Neuro-2a, in which *N-myc* expression is very low and which have been shown to contain a single copy of *N-myc* per haploid genome⁴. Similarly, Southern blot analysis revealed equivalent numbers of *N-myc* copies in genomic DNA isolated from PSA-1, F9, ES and normal mouse spleen cells (data not shown). Therefore, the abundant expression of *N-myc* in EC and ES cells is not the consequence of gene amplification. Similarly, we found that the NTera2 clone D1 human teratocarcinoma cells, which also express relatively high levels of *N-myc* RNA, contain a single *N-myc* copy per haploid genome (data not shown).

In a recent study it was found that when human neuroblastoma cells are induced to differentiate *in vitro* they rapidly undergo a dramatic decrease in *N-myc* expression¹⁷. It seemed

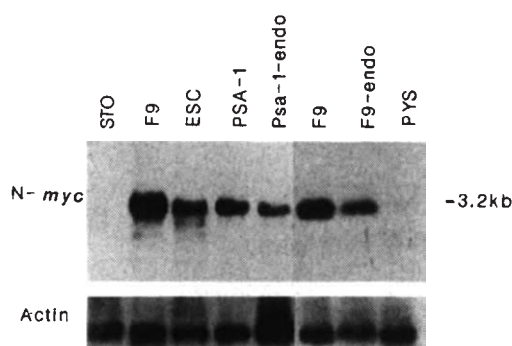


Fig. 1 *N-myc* expression in mouse teratocarcinoma stem cells and their endodermal derivatives.

Methods. STO fibroblasts, F9 embryonal carcinoma cells and PYS endodermal cells were cultured as described previously (refs 9, 11 and 16, respectively). PSA-1 and ES cells (ESC), which are maintained in the undifferentiated state by co-culture with irradiated STO fibroblastic feeders, were cultured and separated from the feeder cells¹³. Endodermal derivatives of the F9 cells (F9-endo) were obtained by incubating cultures of undifferentiated cells with 2×10^{-7} M RA and 10^{-3} M db-cAMP for 4 days¹⁶. Endodermal derivatives of the PSA-1 cells (PSA-1-endo) were obtained by allowing the PSA-1 cells to form embryoid bodies¹⁰. After 6 days of culture in bacteriological dishes, the embryoid bodies were allowed to reattach to a collagen-coated tissue culture substratum; after 2 days, the endodermal cells were collected as described previously²³. Total cytoplasmic RNA was prepared from each of the cell types after lysis of the cell membranes with Nonidet P-40, followed by phenol/chloroform extraction¹⁷. Poly(A)⁺ RNA was isolated by affinity chromatography using oligo(dT)-cellulose and each lane was loaded with ~5 µg of denatured poly(A)⁺ RNA from the designated cell type. The RNAs were separated by electrophoresis in a 1% agarose-formaldehyde gel¹⁷, and the RNA species were transferred to nylon membrane (Gene Screen; NEN), crosslinked to the nylon membrane by ultraviolet irradiation and hybridized with either the nick-translated ³²P-labelled (4 × 10⁸ c.p.m. µg⁻¹) 1-kb *Eco*RI-*Bam*HI fragment of pNb-1 human *N-myc* DNA¹ or chick β-actin DNA (ref. 24; provided by D. Cleveland), in 1 mM EDTA, 0.5 M NaHPO₄ pH 7.2, 7% SDS, and 1% bovine serum albumin (BSA)²⁵. The blot was washed in 1 mM EDTA, 40 mM NaHPO₄ pH 7.2, 1% BSA at 65°C. The Gene Screen membranes were stripped of hybridized probe by washing in 70% formamide, 0.2 × SSC, 0.1% SDS at 65°C for 30 min. Autoradiographic exposure time, with an intensifying screen, was ~16 h for *N-myc* and 4 h for β-actin.

possible that mouse EC cells might show a similar regulation of *N-myc* expression during differentiation to endodermal cells as no *N-myc* expression was detectable in the established endodermal cell line, PYS (Fig. 1). Relatively pure populations of endodermal cells can be derived, by different methods, from the two EC cell lines studied. When PSA-1 cells are cultured as rounded aggregates, the outer cells spontaneously differentiate to endoderm. The two-layered structures that result, consisting of a solid core of undifferentiated cells surrounded by a single layer of endoderm, are known as simple embryoid bodies because of their similarity to the fetal portion of the normal 4.5-day-old mouse embryo^{9,10}. Purified populations of PSA-1-derived endodermal cells were obtained as described in Fig. 1 legend. F9 cells can be induced to differentiate to endoderm at high frequency in monolayer culture by treatment with retinoic acid (RA) and dibutyryl cyclic AMP (db-cAMP)¹⁸. Northern blot analysis of poly(A)⁺ cytoplasmic RNA from PSA-1-derived and F9-derived endodermal cells revealed readily detectable *N-myc* transcripts (Fig. 1). Densitometer tracings of the autoradiograms shown in Fig. 1 indicate that the ratio of *N-myc* RNA to β-actin RNA is approximately two to five times lower in F9-derived and PSA-1-derived endodermal cells than in their undifferentiated progenitors. If the levels of β-actin RNA are the same in EC and endodermal cells, these data suggest that

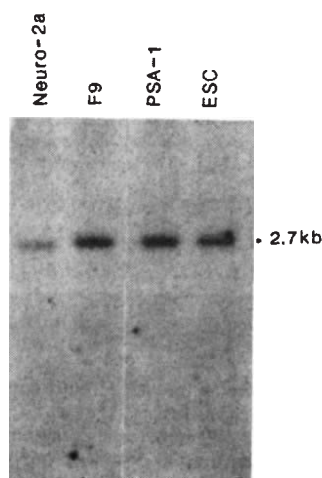


Fig. 2 Southern blot analysis of *N-myc* in genomic DNA from mouse cells. Genomic DNA was isolated from Neuro-2a mouse neuroblastoma cells⁴, PSA-1 and F9 embryonal carcinoma cells and ES cells, and 10 µg of each was digested to completion with *Hind*III. The restriction fragments were separated on an agarose gel and transferred to nitrocellulose, then hybridized with nick-translated ³²P-labelled pNb-1 DNA as described previously¹.

the level of *N-myc* RNA decreases as EC cells differentiate to endoderm.

It is not clear why the 'primary' endodermal cells express readily detectable levels of *N-myc* RNA whereas PYS cells express little or none. It is possible that the *N-myc* expression observed in the endodermal cultures reflects the presence of a subpopulation of undifferentiated cells. However, by morphological criteria the culture of PSA-1-derived cells was almost exclusively of an endodermal phenotype. In addition, when the Northern blot of the RNAs from F9 and F9-endo cells shown in Fig. 1 was rehybridized with a probe for the *c-myc* gene, it was apparent that whereas relatively high levels of *c-myc* RNA are expressed in undifferentiated F9 cells, this RNA is barely detectable in their endodermal derivatives (data not shown). These results, which are in agreement with data from previous studies of *c-myc* expression in F9 cells and their endodermal derivatives¹⁹, suggest that there are relatively few undifferentiated cells in the endodermal populations. Another possible explanation for the differences in *N-myc* expression in primary as compared with PYS endodermal cells is that *N-myc* RNA has a relatively long half-life and that the *N-myc* RNA detected in primary endodermal cells was transcribed before differentiation. Alternatively, it is possible that the primary endodermal cultures contain cells in a state of differentiation which differs from that of PYS cells.

The results of this study raise the question of the functional significance of the abundant *N-myc* gene expression in teratocarcinoma cells. It is conceivable that the relatively high levels of *N-myc* expression in EC cells are a function of their tumorigenicity; alternatively, abundant *N-myc* expression may be a characteristic of early embryonic cells and therefore the *N-myc* expression observed in embryonal carcinoma and embryonic stem cells might be one manifestation of their close similarity to normal early embryonic cells. Although we have not tested this hypothesis directly as it is not feasible to obtain sufficient RNA for Northern blot analysis from the ICM cells to which EC and embryonic stem cells correspond, we have studied *N-myc* expression in embryos at later stages of development. Figure 3 shows the results of Northern blot analysis of *N-myc* expression in embryos collected at daily intervals from 9.5 to 17.5 days of gestation. At each stage the embryo proper was separated from the extraembryonic membranes and placenta, an appropriate number of embryos were pooled, and total cellular poly(A)⁺ RNA was isolated. In addition, total cellular

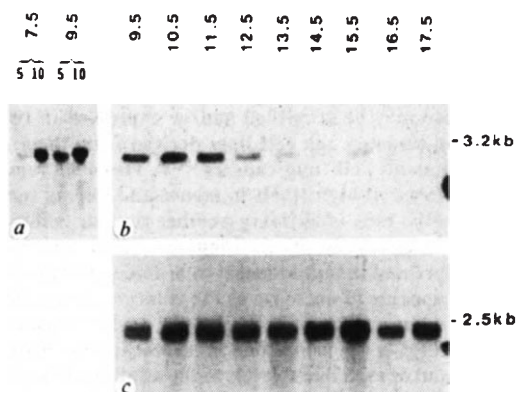


Fig. 3 *N-myc* expression in mouse embryos at various stages of gestation. *a*, Total cellular RNA isolated from embryos at 7.5 and 9.5 days of gestation and hybridized with an *N-myc* probe. The lanes contain ~5 or 10 µg of RNA, as indicated. *b*, Poly(A)⁺ cellular RNA isolated from embryos at days 9.5–17.5 of gestation, hybridized with *N-myc* probe. Each lane contains ~5 µg of RNA. *c*, The blot shown in *b* was stripped of probe as described in Fig. 1 legend and hybridized with a subclone of exon 3 of the human *c-myc* oncogene (*Clal*/*Eco*RI fragment; provided by N. Hay). **Methods.** Embryos were obtained by mating mice of various inbred genotypes. The day on which the genital plug was detected was designated as day 0.5 of gestation. At each stage, total RNA was immediately extracted from a pool of embryos following homogenization in guanidine thiocyanate²⁶. RNA was isolated and analysed as described in Fig. 1 legend.

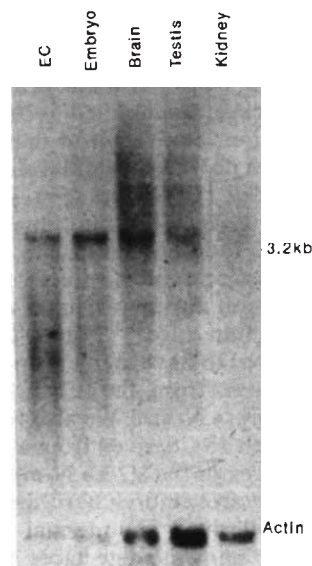


Fig. 4 *N-myc* expression in adult mouse organs. Approximately 4 µg of poly(A)⁺ RNA from PSA-1 teratocarcinoma stem cells (EC) and from embryos at day 11.5 of gestation, and ~10 µg of poly(A)⁺ RNA isolated from organs obtained from randomly bred adult mice, were analysed as described in Fig. 1 legend.

RNA was isolated from a pool of embryos at 7.5 days of gestation. The results of this experiment show that *N-myc* RNA was detectable at all stages tested. Figure 3*a* shows that *N-myc* RNA is detectable in 5 µg of total RNA from embryos at 7.5 and 9.5 days of gestation, indicating that *N-myc* expression is relatively abundant at these stages of development. The level of expression in these embryos is roughly equivalent to that in PSA-1 cells (data not shown). Figure 3*b* shows the level of *N-myc* expression in poly(A)⁺ RNA isolated from embryos at days 9.5–17.5 of gestation. Roughly equal amounts of RNA were loaded in each lane, and this equivalence was confirmed by

rehybridization of the blot with a cloned probe for β -actin (data not shown). From these results it seems that the level of N-myc RNA is highest at days 7.5–11.5 of development and thereafter decreases with increasing time of gestation. The Northern blot shown in Fig. 3b was also rehybridized with a probe for the c-myc gene (Fig. 3c); in contrast to the differential expression observed for N-myc, the levels of c-myc RNA did not change with increasing stage of gestation. The latter results are largely consistent with previous studies²⁰. These data, taken together with the results of studies of the expression of various other oncogenes in mouse embryos^{20–22}, indicate that N-myc, c-myc and other oncogenes are regulated independently during development.

We also tested various organs from adult mice for expression of N-myc. N-myc RNA was readily detected in poly(A)⁺ RNA isolated from brain, was less abundant in RNA from testis and kidney (Fig. 4), but was not detectable in RNA from spleen or liver (data not shown). Densitometer tracings of the autoradiograms indicated that the ratio of N-myc to β -actin RNA is approximately fivefold higher in teratocarcinoma stem cells and in embryos at the mid-gestation stages of development than in brain.

The data presented here demonstrate that the N-myc gene is expressed during normal mouse embryogenesis and in certain organs of the adult. From the results of our study of teratocarcinoma cells, and assuming that these cells provide a valid model system for the study of gene expression in the peri-implantation mouse embryo, it seems likely that N-myc is expressed at relatively high levels by ICM cells at around the time of implantation. In contrast, primary endoderm, which is the first differentiated cell type formed by the ICM, may express it at a lower level. We have demonstrated directly that abundant N-myc expression is characteristic of the mid-gestation embryo and that the level of expression decreases at later stages of embryogenesis. An intriguing possibility is that the abundant N-myc RNA detected in embryos at earlier stages of development is a manifestation of N-myc expression in most or all cells of the embryo, whereas the decreased levels of N-myc RNA at later stages reflect expression in a very limited number of cell types, possibly only in those tissues in which N-myc is expressed in the adult (for example, brain, testis and kidney). The results of *in situ* hybridization studies of embryos and adult organs, now in progress, will help to elucidate the tissue and cell specificity of N-myc expression. The information from such studies should ultimately contribute to our understanding of the normal function of the gene.

This work was supported in part by grants from the American Cancer Society to G.R.M. and from the National Cancer Institute (NIH) to J.M.B., and by funds from the G.W. Hooper Foundation. A.J. is a recipient of a senior postdoctoral fellowship from the California Division of the American Cancer Society. We thank Dr Alexandra Joyner for providing the poly(A)⁺ embryo RNA, Drs Joyner and George Church for technical advice and helpful discussion, and Ms Joy Graesser for technical assistance.

Received 24 June; accepted 19 September 1985.

- Schwab, M. *et al.* *Nature* **305**, 245–248 (1983).
- Kohl, N. E. *et al.* *Cell* **35**, 359–367 (1983).
- Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. *Science* **224**, 1121–1124 (1984).
- Schwab, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 4940–4944 (1984).
- Kohl, N. E., Gee, C. E. & Alt, F. W. *Science* **226**, 1335–1337 (1984).
- Schwab, M., Varmus, H. E. & Bishop, J. M. *Nature* **316**, 160–162 (1985).
- Lee, W. H., Murphree, A. L. & Benedict, W. F. *Nature* **309**, 458–460 (1984).
- Nau, M. M. *et al.* *Curr. Topics Microbiol. Immun.* **113**, 172–177 (1984).
- Martin, G. R. & Evans, M. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1441–1445 (1975).
- Martin, G. R., Wiley, L. M. & Damjanov, I. *Dev. Biol.* **61**, 230–244 (1977).
- Bernstine, E. G., Hooper, M. L., Grandchamp, S. & Ephrussi, B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3899–3903 (1973).
- Martin, G. R. *Science* **209**, 768–776 (1980).
- Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634–7638 (1981).
- Evans, M. J. & Kaufman, M. H. *Nature* **292**, 154–156 (1981).
- Andrews, P. W. *et al.* *Lab. Invest.* **50**, 147–162 (1984).
- Lehman, J. M., Speers, W. C., Schwartzendruber, D. E. & Pierce, G. B. *J. cell. Physiol.* **84**, 13–28 (1974).
- Thiele, C. J., Reynolds, C. P. & Israel, M. A. *Nature* **313**, 404–406 (1985).

- Strickland, S., Smith, K. & Marotti, K. *Cell* **21**, 347–355 (1980).
- Campisi, J., Gray, H. E., Pardee, A. B., Dean, M. & Sonenshein, G. E. *Cell* **36**, 241–247 (1984).
- Slamon, D. J. & Cline, M. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7141–7145 (1984).
- Muller, R., Slamon, D. J., Tremblay, J. M., Cline, M. J. & Verma, I. M. *Nature* **299**, 640–644 (1982).
- Muller, R. *et al.* *Molec. cell. Biol.* **3**, 1062–1069 (1983).
- Grabel, L. B. & Martin, G. R. *Dev. Biol.* **95**, 115–125 (1983).
- Cleveland, D. *et al.* *Cell* **20**, 95–105 (1980).
- Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991–1995 (1984).
- Poole, S. J., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40**, 37–43 (1985).

Isolation, sequence determination and expression in *Escherichia coli* of the isopenicillin N synthetase gene from *Cephalosporium acremonium*

S. M. Samson*, R. Belagaje*, D. T. Blankenship*, J. L. Chapman*, D. Perry†, P. L. Skatrud*, R. M. VanFrank*, E. P. Abraham†, J. E. Baldwin‡, S. W. Queener* & T. D. Ingolia*§

* Eli Lilly & Co., Indianapolis, Indiana 46285, USA

† Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

‡ Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

The enzyme isopenicillin N synthetase (IPS) catalyses the oxidative condensation of δ -(L- α -aminoadipyl)-L-cysteinyl-D-valine (LLD-ACV) to isopenicillin N, which is a central reaction in the pathway to clinically important penicillins and cephalosporins. Here we report the cloning, characterization and expression in *Escherichia coli* of the gene encoding the IPS protein in *Cephalosporium acremonium*. The IPS gene was identified by purifying IPS protein, determining the first 23 amino-terminal amino acids, preparing a set of synthetic oligonucleotides encoding a portion of the determined amino-acid sequence, and probing a cosmid genome library with the mixed oligonucleotides. A cosmid hybridizing with the probe was isolated and the IPS gene was localized and sequenced. The IPS gene encodes a polypeptide of relative molecular mass (M_r) 38,416. When this open reading frame was cloned into an *E. coli* expression vector and inserted into *E. coli*, the recombinant *E. coli* produced a new protein co-migrating with authentic IPS as the major protein of the cell (~20% of cell protein). Crude cell extracts condensed LLD-ACV to a penicillinase-sensitive molecule whose antibacterial activity indicated that it was isopenicillin N.

The conversion of LLD-ACV to isopenicillin N by IPS or cyclase is carried out by a variety of organisms, including *Penicillium chrysogenum*, *C. acremonium*, *Streptomyces clavuligerus* and other Actinomycetes. Penicillins and cephalosporins are isolated from these organisms for the manufacture of important β -lactam antibiotics. The cyclase enzyme, in the presence of Fe²⁺ and ascorbate, removes four hydrogens from LLD-ACV and consumes one oxygen, resulting in the formation of the two rings found in isopenicillin N¹. Most of the studies of the IPS reaction have examined enzymes from *C. acremonium* and *S. clavuligerus*. The *C. acremonium* enzyme has been purified^{2,3}, and used to condense not only LLD-ACV to isopenicillin N, but also variants of LLD-ACV to different β -lactam structures^{4–7}. In order to extend our knowledge of the properties and potential of the IPS enzyme, we have cloned the IPS gene and expressed the activity in *E. coli*.

IPS protein from *C. acremonium* was purified using modified published procedures² with additional reverse-phase HPLC steps. The purified protein was then subjected to amino-terminal sequence analysis, and 23 amino-acid residues were determined (see Fig. 1). Sixty-four synthetic oligonucleotides were prepared, one of which must be the perfect match of the coding information

§ To whom correspondence should be addressed.

Fig. 1 Amino-terminal sequence analysis of IPS and synthesis of oligomeric DNA probes. **A**, The first 23 amino-terminal amino-acid residues of IPS. The question mark indicates that the first cycle of the sequenator run was ambiguous. The underlined sequence (residues 18–23) represents the region encoded by the synthetic oligonucleotides. **B**, The possible RNA codons for the underlined sequence in **A**. As the third position of each codon is variable in this region, all possibilities are shown. **C**, Two sets of 17-mer oligonucleotides incorporating the possibilities shown in **B** were synthesized. Each pool has a complexity of 32. **D**, The experimentally derived DNA sequence and the corresponding predicted amino-acid sequence (see Fig. 3).

Methods. Extraction and purification of IPS from *C. acremonium* were done as described by Pang *et al.*² except that the Dyno-Mill agitator disks were run at 2,000 r.p.m. to keep the temperature below 5°C during disintegration of the mycelium. The peak fractions containing the synthetase were used for further processing. Nearly pure IPS protein was applied to a ProRPC HR 5/10 C₈ reverse-phase column using a Pharmacia FPLC system. The A solvent was 0.1% trifluoroacetic acid and 0.1% morpholine in water; the B solvent was 80% isopropanol, 20% acetonitrile with 0.1 ml per 100 ml each of trifluoroacetic acid and morpholine. The column was run at 0.3 ml min⁻¹ with a series of the following linear gradients: from 0 to 5 ml of elution buffer, B went from 0 to 30%; from 5 to 23 ml B went from 30 to 70%; and from 23 to 25 ml B went from 70 to 100%. The 64 deoxyoligonucleotides were chemically synthesized in two pools of 32 17-mers each, using an Applied Biosystem DNA Synthesizer (Model 380-A) according to the manufacturer's recommended protocols. Oligomers were purified by 20% polyacrylamide/7 M urea gel electrophoresis.

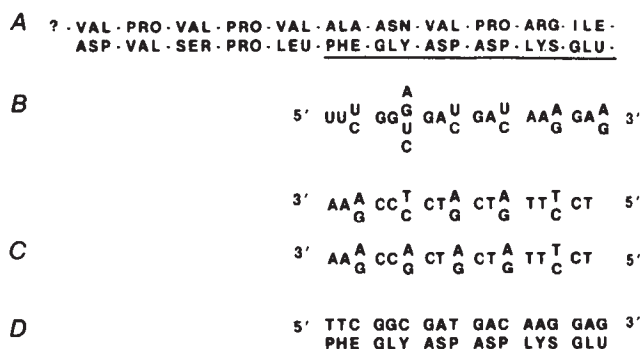
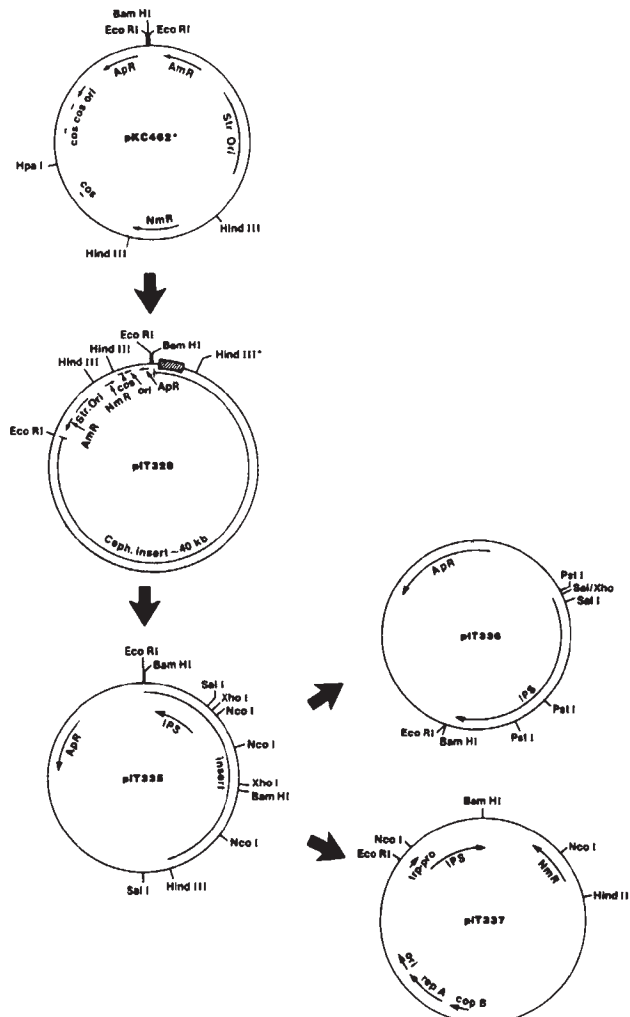


Fig. 2 Schematic diagrams of cosmids and plasmids used in the present study. The size of the circle is not proportional to the size of the plasmid. Except where indicated below, if a restriction enzyme cut site is indicated on the drawing, all restriction sites for that enzyme are included in the diagram. pKC462*, a 12.4-kb cosmid constructed by R. N. Rao and R. Stanzak (personal communication) which contains a *Streptomyces* origin of replication (Str Ori) from pFJ103 (ref. 10), and an *E. coli* origin of replication (ori) from pBR322. Antibiotic resistance markers on this plasmid are the ampicillin resistance gene (ApR) from pBR322, the apramycin resistance gene (AmR) from pKC222 (ref. 11), and the neomycin resistance gene (NmR) from Tn5. The three cos sites (cos) are from bacteriophage λ . pKC462* contains a unique *Bam*HI restriction site (used in this study for cloning) flanked by unique *Eco*RI restriction sites. pIT328, an ~50-kb plasmid created by ligating *C. acremonium* genomic DNA partially cleaved with restriction enzyme *Mbo*I into the *Bam*HI site of pKC462*. pIT328 contains ~40 kb of *C. acremonium* DNA (Ceph. insert). The cross-hatched box indicates the protein-coding region of the IPS gene. *Hind*III* is one of at least 10 *Hind*III sites in the insert, the rest of which are not shown. pIT335, an 8.2-kb plasmid, was constructed by digesting pIT328 with *Hind*III and re-ligating, thus eliminating half the vector and all but ~4.5 kb of the *Cephalosporium* DNA insert. The IPS gene is transcribed in the direction of the arrow. Only the relevant restriction sites are shown, and one more unmapped *Nco*I restriction site exists outside the IPS gene in pIT335. pIT336, a 4.1-kb plasmid, was created by cloning a 1,400-bp *Eco*RI-*Xho*I fragment from pIT335 into the *Eco*RI-*Sal*I sites of pUC8 (ref. 12). The *Cephalosporium* DNA insert (IPS) contains most of the protein-coding region of the IPS gene. pIT337 is an 11.6-kb expression plasmid containing the complete coding region of the IPS gene. The 1,460-bp *Nco*I-*Bam*HI fragment from pIT335 was ligated with two fragments (an 8.7-kb *Nco*I-*Nco*I fragment and a 1.5-kb *Nco*I-*Bam*HI fragment) from pCZ106. pCZ106 is a derivative of pCZ115 (ref. 13) which contains the bovine growth hormone gene transcribed from the *E. coli* tryptophan promoter and a temperature-sensitive copy control (runaway replicon) system. The *E. coli* tryptophan promoter (*trp* pro) drives transcription of the IPS gene. The location and direction of transcription of the IPS gene and of the neomycin resistance gene (NmR) are also indicated.

Methods. *C. acremonium* genomic DNA was isolated from protoplasts prepared according to Queener *et al.*¹⁴. The protoplasts were lysed by incubating them in 3% *N*-lauroyl sarkosyl, 0.5 M Tris-HCl pH 9.0, 0.2 M EDTA at 65°C for 15 min. The mixture was cooled and then extracted with buffered phenol (10 min) and chloroform (10 min). The DNA layer was mixed with a 1:10 volume of 5 M KOAc and spooled with 2 vol. of cold 100% ethanol. DNA was dissolved in 2 ml TE buffer (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA) then treated with 50 μ g ml⁻¹ proteinase K (Sigma; Type XI) at 65°C for 1 h in 0.01 M Tris-HCl pH 7.8, 5 mM EDTA, 0.5% SDS. The solution was then extracted for 10 min each with buffered phenol and chloroform and the DNA respoiled and suspended in 0.5 ml TE buffer. Partial enzymatic digestions with *Mbo*I were done to achieve an average size of 30–40 kb. pKC462* was digested with *Hpa*I, treated with alkaline phosphatase, then cut with *Bam*HI. The partially digested genomic DNA was ligated into the *Bam*HI site of pKC462* and the recombinant DNA molecules were packaged into phage λ heads using Promega Biotec packaging extracts according to the recommendations of the manufacturer. *E. coli* SF8 cells grown in the presence of 0.2% maltose were transduced with the recombinant λ particles and cells containing cosmids were selected on plates containing 100 μ g ml⁻¹ ampicillin and then grown in liquid culture in microtitre dishes in the presence of 100 μ g ml⁻¹ apramycin. The library was screened according to the method of Grunstein and Hogness¹⁵. Oligonucleotide probes (described in Fig. 1 legend) were kinased under conditions described previously¹⁶. The nitrocellulose filters containing the recombinant *E. coli* were hybridized at 42°C for 2 h in 10 $\times$ Denhardt's, 6 $\times$ SSC, 0.1% sodium pyrophosphate¹⁷. Washes were at room temperature in 6 $\times$ SSC for 30 min, followed by another 5-min wash in 6 $\times$ SSC at 42–58°C.



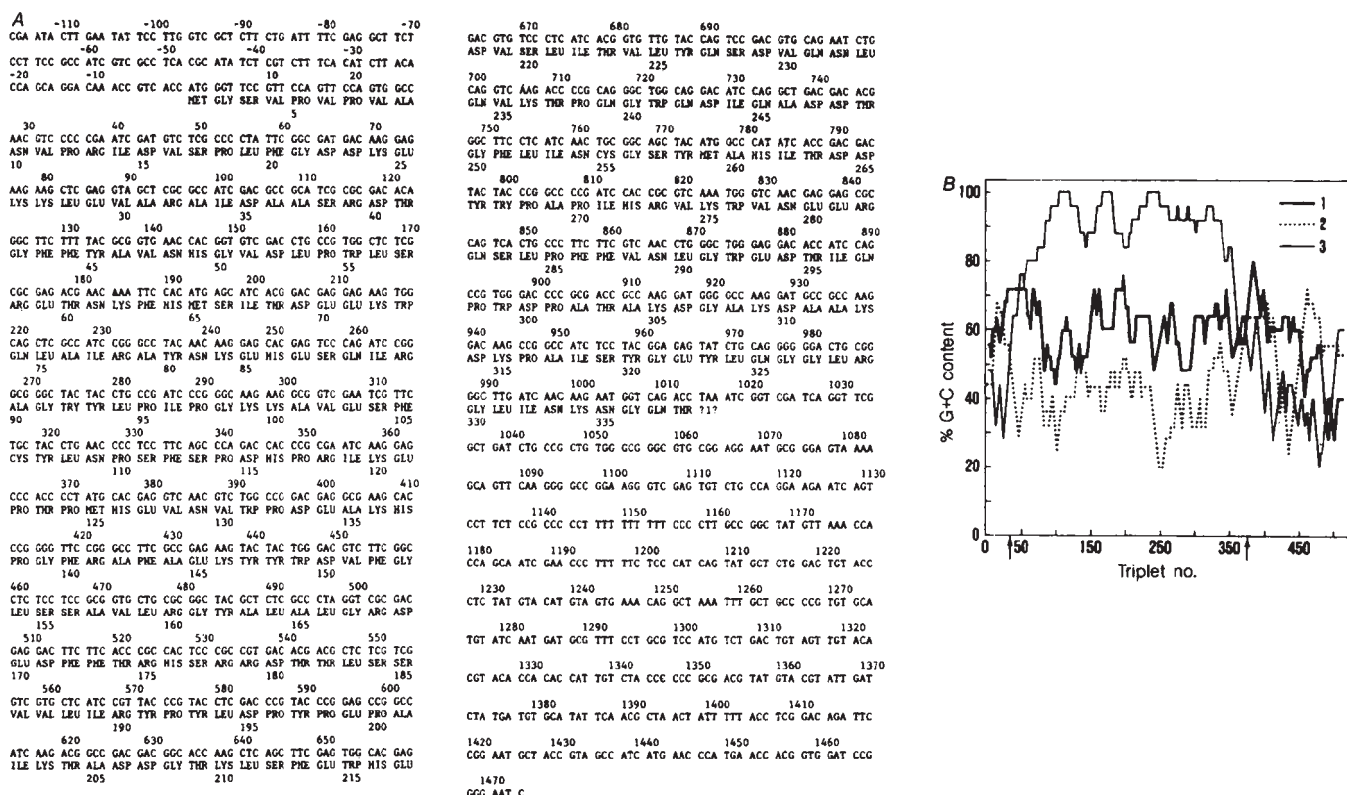


Fig. 3 **A**, Complete nucleotide and predicted amino-acid sequence of isopenicillin N synthetase gene from pIT335. The DNA sequence was determined by the chemical method of Maxam and Gilbert⁸ as described previously¹⁶. Both strands were sequenced for over 95% of the protein-coding region, and the remaining areas were sequenced at least twice from different sites on the same strand. The primary sequence of the IPS gene, together with 118 bp preceding the initiation codon and 455 bp after the termination codon, is shown with the predicted amino-acid sequence of a 1,017-bp open reading frame. Numbers above the nucleotide sequence refer to base-pair location. Amino-acid residues are numbered below the protein sequence. The proposed translation initiation site is amino-acid 1. **B**, Analysis of G+C content as a function of position in the codon. Each curve represents the G+C content of one of the three codon positions. The per cent G+C at each base is calculated as a forward average of 20 codons according to the algorithm of Bibb *et al.*⁹. For clarity, the curves presented are smoothed slightly relative to the original computer print-out. The heavy line corresponds to the first base of a codon, the dotted line refers to the second base, and the thin line refers to the third position. The arrows along the axis indicate the limits of the IPS protein-coding region (see **A**).

for amino acids 18–23 (Phe-Gly-Asp-Asp-Lys-Glu) encoded by the IPS gene. The 64 possible 17-mer oligonucleotides were synthesized in two pools of 32 (Fig. 1). The pool of 32 that hybridized most efficiently to restriction enzyme-cleaved *C. acremonium* genomic DNA immobilized on nitrocellulose was used for further experiments.

The synthetic 17-mer oligonucleotides were used to screen a cosmid genome bank for the IPS gene. The region of one cosmid, pIT328, which hybridized to the oligonucleotide probes was determined to be adjacent to the *E. coli* selectable markers by restriction mapping and Southern hybridizations, so the cosmid was truncated by restriction enzyme *Hind*III digestion and re-ligation to yield the 11-kilobase (kb) plasmid pIT335 (Fig. 2). The DNA sequence of a portion of pIT335 was determined by the Maxam-Gilbert technique⁸ and the predicted amino-acid sequence of one region matched the experimentally determined amino-acid sequence of the IPS protein (Fig. 3). The open reading frame containing this information (Fig. 3) can encode a polypeptide of M_r 38,416. The reported M_r of the IPS protein, as determined by SDS-polyacrylamide gel electrophoresis, is 40,000 (ref. 3) (see also Fig. 4).

The predicted amino-acid sequence of the IPS protein from pIT337 (Fig. 3) begins with methionine and glycine residues which are not found in the protein isolated from *Cephalosporium* cells. These residues are evidently cleaved post-translationally in *Cephalosporium*.

The protein-coding region of the IPS gene consisted of 63% G+C base pairs. Because of the high G+C content, we used an algorithm described by Bibb *et al.*⁹ to examine the G+C content of the DNA sequence as a function of position in the triplet codons. Bibb *et al.* have shown that the G+C content of the

positions in the triplet codons is biased within protein-coding regions but relatively random outside protein-coding regions for organisms with DNA of high G+C content. In particular, Bibb *et al.* found that the G+C content of the first position in the codon was about average, while that of the second position was much lower, and that of the third position much higher than average. As shown in Fig. 3, the same pattern of G+C content bias was observed for the open reading frame encoding the IPS protein. This provides further evidence that the open reading frame identified by sequencing is correct for its entire length, since a shift in reading frame would have also shifted the G+C content bias.

To further prove that the open reading frame encodes the IPS protein, we fused the open reading frame to an *E. coli* expression system. The open reading frame from pIT335 was isolated on an *Nco*I–*Bam*HI fragment. The recognition sequence for the restriction enzyme *Nco*I, 5'-CCATGG, surrounds the methionine codon that initiates translation of the open reading frame (Fig. 3). The restriction enzyme *Bam*HI cleaves ~450 base pairs (bp) downstream from the proposed translation termination site. The *Nco*I–*Bam*HI fragment was ligated into an *E. coli* expression vector to create pIT337. The *E. coli* expression vector used the promoter and ribosome binding site from the *E. coli* tryptophan operon and a temperature-sensitive copy control system ('runaway replicon'). *E. coli* cells transformed with pIT337 produce a new protein which co-migrates with authentic IPS protein on SDS-polyacrylamide gels. At low temperatures the cells produce low amounts of the new protein, but at high temperatures they produce about 20% of total cell protein as the putative IPS polypeptide (Fig. 4). The other new major protein band produced by the transformed *E. coli* cells is

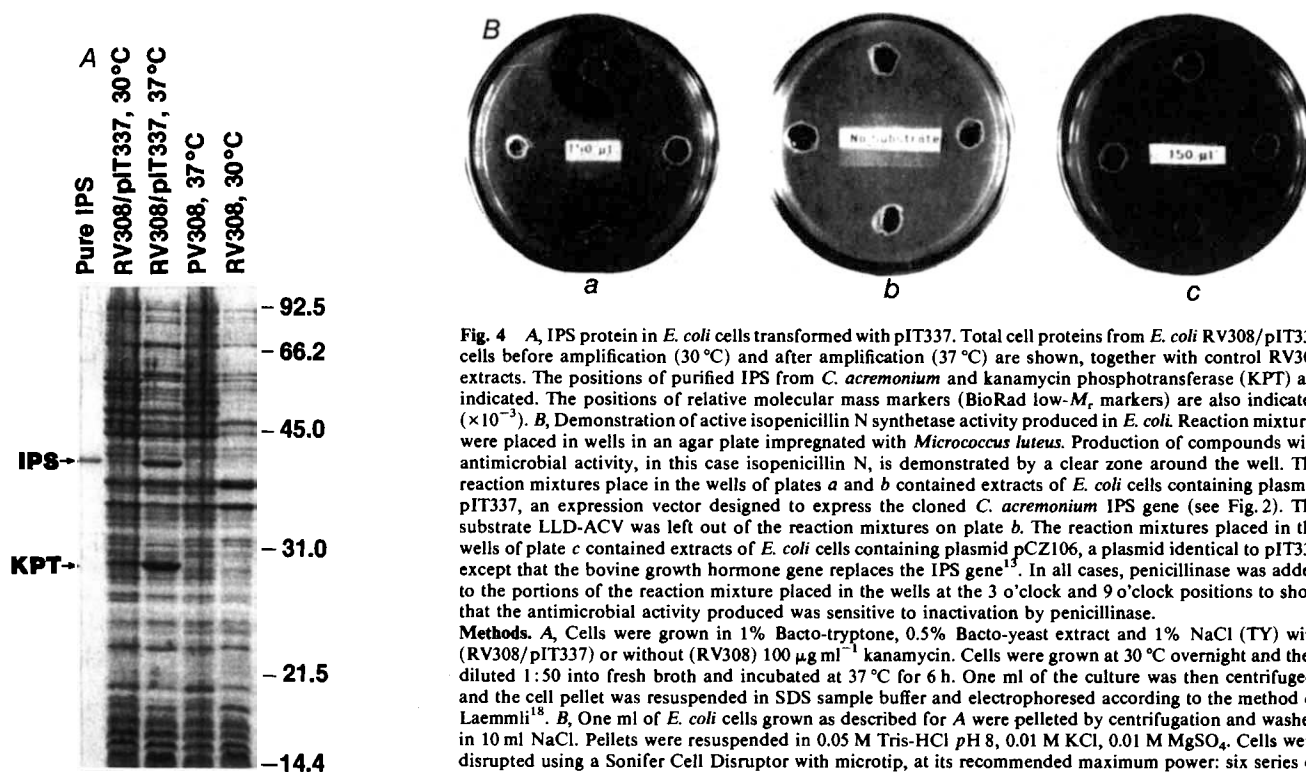


Fig. 4 **A**, IPS protein in *E. coli* cells transformed with pIT337. Total cell proteins from *E. coli* RV308/pIT337 cells before amplification (30 °C) and after amplification (37 °C) are shown, together with control RV308 extracts. The positions of purified IPS from *C. acremonium* and kanamycin phosphotransferase (KPT) are indicated. The positions of relative molecular mass markers (BioRad low- M_r markers) are also indicated ($\times 10^{-3}$). **B**, Demonstration of active isopenicillin N synthetase activity produced in *E. coli*. Reaction mixtures were placed in wells in an agar plate impregnated with *Micrococcus luteus*. Production of compounds with antimicrobial activity, in this case isopenicillin N, is demonstrated by a clear zone around the well. The reaction mixtures placed in the wells of plates **a** and **b** contained extracts of *E. coli* cells containing plasmid pIT337, an expression vector designed to express the cloned *C. acremonium* IPS gene (see Fig. 2). The substrate LLD-ACV was left out of the reaction mixtures on plate **b**. The reaction mixtures placed in the wells of plate **c** contained extracts of *E. coli* cells containing plasmid pCZ106, a plasmid identical to pIT337 except that the bovine growth hormone gene replaces the IPS gene¹³. In all cases, penicillinase was added to the portions of the reaction mixture placed in the wells at the 3 o'clock and 9 o'clock positions to show that the antimicrobial activity produced was sensitive to inactivation by penicillinase.

Methods. **A**, Cells were grown in 1% Bacto-tryptone, 0.5% Bacto-yeast extract and 1% NaCl (TY) with (RV308/pIT337) or without (RV308) 100 $\mu\text{g ml}^{-1}$ kanamycin. Cells were grown at 30 °C overnight and then diluted 1:50 into fresh broth and incubated at 37 °C for 6 h. One ml of the culture was then centrifuged, and the cell pellet was resuspended in SDS sample buffer and electrophoresed according to the method of Laemmli¹⁸. **B**, One ml of *E. coli* cells grown as described for **A** were pelleted by centrifugation and washed in 10 ml NaCl. Pellets were resuspended in 0.05 M Tris-HCl pH 8, 0.01 M KCl, 0.01 M MgSO_4 . Cells were disrupted using a Sonifier Cell Disruptor with microtip, at its recommended maximum power: six series of 5-s bursts with 60-s intervals were conducted. Debris was centrifuged out and aliquots (25, 50, 100 and 150 μl) were assayed for isopenicillin N synthetase activity according to Shen *et al.*¹⁹.

kanamycin phosphotransferase, which is also encoded on the *E. coli* expression vector.

Crude cell-free extracts from cells containing pIT337 also contain IPS activity. Extracts were incubated with LLD-ACV and the reaction mixture tested for antimicrobial activity on agar plates impregnated with *Micrococcus luteus*. As shown in Fig. 4, generation of antimicrobial activity in the assay was dependent on the presence of the IPS gene in the *E. coli* expression vector and on exogenous addition of LLD-ACV. These results indicate that the open reading frame from pIT335 encodes a polypeptide with IPS activity, and that the polypeptide is expressed in an active form in *E. coli*.

The cloning and expression of the IPS gene from *C. acremonium* will facilitate further analysis of the penicillin-cephalosporin biosynthetic pathway. The enzymatic properties and structure/function relationships of IPS can now be tested

in more detail because of the greater availability of the synthetase. The properties of the *Cephalosporium* gene, especially with respect to the control of transcription, can also be studied. Control of expression of the IPS gene is particularly interesting because β -lactam antibiotics are usually produced late in the growth cycle, suggesting the possibility of control of expression of the biosynthetic genes themselves. These aspects are presently under investigation.

We thank the Lilly Research Laboratories (a division of Eli Lilly & Co.) for financial support and the following individuals for encouragement: R. H. Baltz, J. P. Burnett, N. Neuss, E. Flynn, J. Whitney, R. Losick and D. Dennen. We also thank R. N. Rao for providing pKC462*; B. Weigel, R. Stanzak, G. Wiseman, S. Burgett, J. Shrote, C. Beckage and G. Griesinger for advice and assistance; and R. H. Baltz for a critical review of the manuscript.

Received 3 June; accepted 10 September 1985.

1. Abraham, E. P., Huddleston, J. A., Jayatilake, G. S., O'Sullivan, J. & White, R. L. *Spec. Publ. chem. Soc.* **38**, 125-134 (1981).
2. Pang, C.-C. *et al. Biochem. J.* **222**, 789-795 (1984).
3. Hollander, I. J., Shen, Y.-Q., Heim, J., Demain, A. L. & Wolfe, S. *Science* **224**, 610-612 (1984).
4. Bahadur, G. A. *et al. J. Am. chem. Soc.* **103**, 7650-7651 (1981).
5. Baldwin, J. E. *et al. JCS chem. Commun.*, 1317-1319 (1983).
6. Baldwin, J. E. *et al. JCS chem. Commun.*, 1225-1227 (1984).
7. Wolfe, S., Hollander, I. J. & Demain, A. L. *Biotechnology* **2**, 635-636 (1984).
8. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560 (1977).
9. Bibb, M. J., Findlay, P. R. & Johnson, M. W. *Gene* **30**, 157-166 (1984).

10. Richardson, M. A., Mabe, J. A., Beerman, N. E., Nakatsukasa, W. M. & Fayerman, J. T. *Gene* **20**, 451 (1982).
11. Rao, R. N. *et al. Antimicrob. Ag. Chemother.* **24**, 689-695 (1983).
12. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
13. Schoner, B. E., Hsiung, H. M., Belagaje, R. M., Mayne, N. G. & Schoner, R. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5403-5407 (1984).
14. Queener, S. W., Ingolia, T. D., Skatrud, P. L., Chapman, J. L. & Kaster, K. R. in *Microbiology 1985* (ed. Leive, L.) (American Society of Microbiology, in the press).
15. Grunstein, M. & Hogness, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
16. Ingolia, T. D. & Craig, E. A. *Nucleic Acids Res.* **9**, 1627-1642 (1981).
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 326-328 (Cold Spring Harbor Laboratory, New York, 1982).
18. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
19. Shen, Y.-Q., Wolfe, S. & Demain, A. L. *J. Antibiot.* **37**, 1044-1048 (1984).